



TUGAS AKHIR – SS141501

**PENGELOMPOKAN KABUPATEN/KOTA DI JAWA
TIMUR BERDASARKAN INDIKATOR
PEMBANGUNAN EKONOMI MENGGUNAKAN
MODEL-BASED CLUSTERING**

**SHIDDIQUL AKHYAR
NRP 1313 100 068**

**Dosen Pembimbing
Raden Mohamad Atok, Ph.D
Dr.rer.pol. Dedy Dwi Prastyo**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



TUGAS AKHIR – SS141501

**PENGELOMPOKAN KABUPATEN / KOTA DI JAWA
TIMUR BERDASARKAN INDIKATOR
PEMBANGUNAN EKONOMI MENGGUNAKAN
MODEL-BASED CLUSTERING**

**SHIDDIQUL AKHYAR
NRP 1313 100 068**

**Dosen Pembimbing
Raden Mohamad Atok, Ph.D
Dr.rer.pol. Dedy Dwi Prastyo**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



FINAL PROJECT – SS 141501

**THE CLUSTER OF REGENCY / CITY IN EAST
JAVA BASED ON ECONOMIC DEVELOPMENT
INDICATORS BY USING MODEL-BASED
CLUSTERING ANALYSIS**

**SHIDDIQUL AKHYAR
NRP 1313 100 068**

**Supervisor
Raden Mohamad Atok, Ph.D
Dr.rer.pol. Dedy Dwi Prastyo**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**

LEMBAR PENGESAHAN

PENGELOMPOKAN KABUPATEN / KOTA DI JAWA TIMUR BERDASARKAN INDIKATOR PEMBANGUNAN EKONOMI MENGGUNAKAN MODEL-BASED CLUSTERING

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada
Program Studi Sarjana Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :

Shiddiqul Akhyar
NRP. 1313 100 083

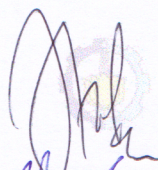

Disetujui oleh Pembimbing Tugas Akhir:

R. Mohamad Atok, Ph.D

NIP: 19710915 199702 1 001

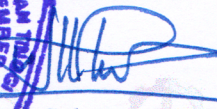
Dr. rer. pol. Dedy Dwi Prastyo

NIP: 19831204 200812 1 002

()
()



Mengetahui,
Kepala Departemen


Dr. Suhartono

NIP. 19710929 199512 1 001

SURABAYA, JULI 2017

PENGELOMPOKAN KABUPATEN / KOTA DI JAWA TIMUR BERDASARKAN INDIKATOR PEMBANGUNAN EKONOMI MENGGUNAKAN MODEL-BASED CLUSTERING

Nama Mahasiswa : Shiddiqul Akhyar
NRP : 1313100068
Departemen : Statistika
Dosen Pembimbing 1 : Raden Mohamad Atok, Ph.D
Dosen Pembimbing 2 : Dr.rer.pol. Dedy Dwi Prastyo

Abstrak

Pembangunan ekonomi merupakan proses untuk meningkatkan pendapatan perkapita selama kurun waktu tertentu. Dalam pembangunan berlingkup luas diperlukan kerjasama antar daerah sangat untuk mempercepat realisasi pembangunan ekonomi. Ketimpangan pembangunan ekonomi terjadi akibat adanya kesenjangan pembangunan. Indikator keberhasilan pembangunan ekonomi didasarkan pada tiga indikator yaitu Indikator moneter, indikator non moneter dan indikator campuran. Meskipun perekonomian Jawa Timur pada tahun 2016 tumbuh sebesar 5,55 persen dari tahun sebelumnya, namun analisis ketimpangan ekonomi masih menunjukkan hasil yang tinggi. Pengelompokan Kabupaten/Kota diperlukan untuk menentukan kebijakan yang didasarkan pada indikator pembangunan ekonomi. Metode model-based clustering menawarkan hasil pengelompokan yang lebih baik. Hasil penelitian menunjukkan kelompok Kabupaten/Kota besar dengan perekonomian maju cenderung menjadi minoritas. Sedangkan kelompok dengan anggota yang konsisten berada pada klaster dengan karakteristik perekonomian menengah ke bawah.

Kata Kunci : Jawa Timur, Ketimpangan, Model-Based Clustering, pembangunan ekonomi

(halaman ini sengaja dikosongkan)

THE CLUSTER OF REGENCY / CITY IN EAST JAVA BASED ON ECONOMIC DEVELOPMENT INDICATORS BY USING MODEL-BASED CLUSTERING ANALYSIS

Name of Student : Shiddiqul Akhyar
NRP : 1313100068
Department : Statistics
Supervisor 1 : Raden Mohamad Atok, Ph.D
Supervisor 2 : Dr.rer.pol. Dedy Dwi Prastyo

Abstract

Economic development can be interpreted as process to improve percapita income during period of time. Good cooperation became a necessary in term of macro economic development. Economic development can be achieved if the government's cooperation among regions are integrative, synchronization, and synergic. There will be a gap in economic development if there is a gap among regions. The succesfulness of economic development can be seen from monetary, non monetary, and both two indicators. In 2016 economic development index in East Java was 5.55%. However, East Java also has high gap in economic index. According to Williamson index, East Java Province still has high index. It means there is wide gap among regions. The specific and directed policy must be implemented to those regions to minimize or reduce economic development gap. Model-based clustering is one of approachs to solve the gap. The model-based clustering method offers better grouping results. The Result of research indicate group of big Regency/City with developed economics tend to become minority. Beside that, the groups with consistent members are in clusters with characteristics of the middle to lower economy.

Keywords : East Java, Gap, Model-Based Clustering, Economic development.

(halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat, hidayah serta karunia-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **“Pengelompokan Kabupaten/ Kota di Jawa Timur berdasarkan Indikator Pembangunan Ekonomi menggunakan Model-Based Clustering”**

Tugas Akhir dapat terselesaikan dengan baik dan lancar tidak lepas dari banyaknya bantuan yang diberikan oleh berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada :

1. Raden Mohamad Atok, Ph.D dan Dr.rer.pol. Dedy Dwi Prastyo selaku dosen pembimbing yang telah meluangkan waktu untuk membimbing pengerjaan Tugas Akhir.
2. Dr. Suhartono selaku Kepala Departemen Statistika ITS yang telah menyediakan fasilitas untuk menyelesaikan Tugas Akhir.
3. Prof.Dr. I Nyoman Budiantara, M.Si dan Imam Safawi Ahmad, M.Si selaku dosen penguji yang telah memberikan saran-saran untuk kesempurnaan Tugas Akhir ini.
4. Dr. Sutikno, selaku Ketua Prodi S1 Departemen Statistika ITS dan segenap dosen yang telah mendidik selama perkuliahan.
5. Dr. Irhamah, selaku dosen wali yang telah memberikan nasehat, motivasi, serta bimbingan selama perkuliahan.
6. Kedua orang tua dan keluarga atas doa, nasehat, kasih sayang, dan dukungan untuk menyelesaikan Tugas Akhir ini.
7. Semua pihak yang telah memberikan dukungan yang tidak dapat penulis sebutkan satu persatu.

Pembuatan laporan Tugas Akhir ini masih jauh dari kesempurnaan, besar harapan dari penulis untuk menerima kritik dan saran yang berguna untuk perbaikan di masa mendatang. Semoga laporan Tugas Akhir ini bermanfaat.

Surabaya, Juli 2017
Penulis

(Halaman ini sengaja dikosongkan)

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
TITLE PAGE	iii
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan	5
1.4 Manfaat	5
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	
2.1 Statistika Deskriptif	7
2.2 Analisis <i>Cluster</i>	7
2.3 Distribusi <i>t</i> Multivariat	10
2.4 Deteksi <i>Outlier</i> Multivariat.....	11
2.5 <i>Model-Based Clustering</i>	12
2.6 Model <i>Finite Mixture</i>	14
2.7 Algoritma <i>Expectation Maximization</i>	15
2.8 <i>Integrated Completed Likelihood</i>	17
2.9 Uji Manova	17
2.10 Pembangunan Ekonomi	18
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data	25

3.2	Variabel Penelitian	25
3.3	Langkah Penelitian	26
3.4	Diagram Alir	27
BAB IV ANALISIS DAN PEMBAHASAN		
4.1	Deteksi <i>Outlier</i> Multivariat.....	29
4.2	Pengelompokan Kabupaten/Kota Kriteria ICL.....	31
4.2.1	Pengelompokan Kabupaten/Kota Tahun 2011	32
4.2.2	Pengelompokan Kabupaten/Kota Tahun 2012.....	34
4.2.3	Pengelompokan Kabupaten/Kota Tahun 2013.....	36
4.2.4	Pengelompokan Kabupaten/Kota Tahun 2014.....	38
4.2.5	Pengelompokan Kabupaten/Kota Tahun 2015.....	40
4.3	Uji Kesamaan Kelompok	42
4.4	Analisis Karakteristik Kelompok	44
BAB V KESIMPULAN DAN SARAN		
5.1	Kesimpulan	49
5.2	Saran	49
DAFTAR PUSTAKA		51
LAMPIRAN		55
BIODATA PENULIS		

DAFTAR GAMBAR

	Halaman
Gambar 2.1	Ilustrasi Bentuk <i>Cluster</i> Berdasarkan Matriks Varians Kovarians 13
Gambar 3.1	Diagram Alir..... 27
Gambar 4.1	<i>Dd Plot</i> Data Tahun 2011 30
Gambar 4.2	<i>Marginal Contour Plot</i> Data Tahun 2011..... 33
Gambar 4.3	<i>Marginal Contour Plot</i> Data Tahun 2012..... 35
Gambar 4.4	<i>Marginal Contour Plot</i> Data Tahun 2013..... 37
Gambar 4.5	<i>Marginal Contour Plot</i> Data Tahun 2014..... 39
Gambar 4.6	<i>Marginal Contour Plot</i> Data Tahun 2015..... 41
Gambar 4.7	Peta Sebaran Pengelompokkan Tahun 2011.... 44
Gambar 4.8	Peta Sebaran Pengelompokkan Tahun 2012.... 45
Gambar 4.9	Peta Sebaran Pengelompokkan Tahun 2013.... 45
Gambar 4.10	Peta Sebaran Pengelompokkan Tahun 2014.... 47
Gambar 4.11	Peta Sebaran Pengelompokkan Tahun 2015.... 47

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

	Halaman
Tabel 2.1 Struktur Data Kelompok ke- <i>g</i> pada Model <i>Finite Mixture</i>	14
Tabel 3.1 Variabel Penelitian	25
Tabel 3.2 Struktur Data dalam Penelitian.....	26
Tabel 4.1 Daftar Kabupaten/Kota yang terdeteksi sebagai data outlier	31
Tabel 4.2 Hasil pengelompokan Kabupaten/Kota di Jawa Timur tahun 2011	34
Tabel 4.3 Hasil pengelompokan Kabupaten/Kota di Jawa Timur tahun 2012	36
Tabel 4.4 Hasil pengelompokan Kabupaten/Kota di Jawa Timur tahun 2013	38
Tabel 4.5 Hasil pengelompokan Kabupaten/Kota di Jawa Timur tahun 2014	40
Tabel 4.6 Hasil pengelompokan Kabupaten/Kota di Jawa Timur tahun 2015	42
Tabel 4.7 Hasil Uji Kesamaan Kelompok dengan nilai <i>Wald Type</i>	43
Tabel 4.8 Hasil Uji Kesamaan Kelompok secara parsial	43
Tabel 4.9 Nilai <i>Mean</i> pada Hasil Pengelompokkan Tahun 2011-2015.....	48

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data Indikator Pembangunan Ekonomi Tahun 2011	55
Lampiran 2 Data Indikator Pembangunan Ekonomi Tahun 2012	57
Lampiran 3 Data Indikator Pembangunan Ekonomi Tahun 2013	59
Lampiran 4 Data Indikator Pembangunan Ekonomi Tahun 2014	61
Lampiran 5 Data Indikator Pembangunan Ekonomi Tahun 2015	63
Lampiran 6 Output Deteksi <i>Outlier</i> data tahun 2011	65
Lampiran 7 Output Deteksi <i>Outlier</i> data tahun 2012	65
Lampiran 8 Output Deteksi <i>Outlier</i> data tahun 2013	66
Lampiran 9 Output Deteksi <i>Outlier</i> data tahun 2014	66
Lampiran 10 Output Deteksi <i>Outlier</i> data tahun 2015	67
Lampiran 11 Tata Nama <i>Model-Based Clustering</i> dalam <i>Software R Package Teigen</i>	68
Lampiran 12 Output nilai ICL data tahun 2011	69
Lampiran 13 Output nilai ICL data tahun 2012	69
Lampiran 14 Output nilai ICL data tahun 2013	70
Lampiran 15 Output nilai ICL data tahun 2014	70
Lampiran 16 Output nilai ICL data tahun 2015	71
Lampiran 17 Output klasifikasi	72
Lampiran 18 Output perbandingan <i>mean</i>	73
Lampiran 19 Surat Legalitas Data	75

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Tingkat perekonomian merupakan salah satu faktor yang seringkali digunakan untuk menyatakan keberhasilan sebuah wilayah atau pemerintahan dalam mengelola kesejahteraan dan kemakmuran masyarakat. Sejahtera dapat diartikan sebagai kondisi masyarakat yang berkecukupan dan mampu memenuhi kebutuhan, sedangkan kemakmuran merupakan bagian yang memungkinkan orang-orang bermasyarakat dengan baik, tenang dan tidak menimbulkan kecemburuan sosial (Dumairy, 1996). Secara sudut pandang ekonomi, hal tersebut dapat dikatakan tercapai ketika Produk Domestik Bruto (PDB) sudah melampaui tingkat pertumbuhan penduduk. Salah satu cara dalam mewujudkan tujuan tersebut adalah melakukan pembangunan pada sektor ekonomi. Pembangunan ekonomi merupakan suatu proses untuk meningkatkan pendapatan per kapita selama kurun waktu yang panjang, dengan catatan bahwa jumlah penduduk yang hidup di bawah garis kemiskinan absolut tidak meningkat dan distribusi pendapatan tidak semakin timpang (Kuncoro, 2006). Proses tersebut memerlukan suatu kenyataan fisik dan tekad suatu masyarakat untuk berupaya sekeras mungkin melalui serangkaian kombinasi proses sosial, ekonomi dan institusional (Todaro & Smith, 2006). Dalam lingkup yang lebih kecil pemerintah daerah dan masyarakat juga bekerjasama mengelola sumber daya yang ada untuk menciptakan suatu lapangan kerja baru dan merangsang perkembangan kegiatan ekonomi dalam wilayah tersebut melalui pembangunan ekonomi daerah (Arsyad, 1999).

Kerjasama antar daerah merupakan faktor yang sangat diperlukan dalam pembangunan ekonomi yang berlingkup luas. Integrasi, sinkronisasi dan sinergi antar daerah sangat penting

untuk mempercepat realisasi pembangunan di daerah (Syafrizal, 2008). Namun tujuan tersebut tidak mudah dicapai karena setiap daerah memiliki perbedaan potensi sumber daya alam, sumber daya manusia dan lembaga institusi. Pemerintahan pada setiap daerah pasti memiliki sudut pandang berbeda dalam mengelola potensi daerah. Terlebih dengan diberlakukannya UU RI nomor 9 tahun 2015 tentang peranan pemerintah untuk lebih dominan dalam menentukan kebijakan di daerah menyebabkan setiap daerah cenderung bersaing menaikkan perekonomian secara regional. Pertumbuhan ekonomi yang cepat akan menimbulkan ketimpangan distribusi pendapatan jika besar atau kecilnya tingkat pertumbuhan penduduk atau perubahan struktur ekonomi tidak ikut diperhatikan. Ketimpangan pembangunan ekonomi secara wajar memang akan terjadi dengan adanya perbedaan tersebut, karena pada dasarnya kesenjangan pembangunan erat kaitannya dengan proses pembangunan itu sendiri (Ardani, 1992). Akibatnya pertumbuhan ekonomi yang diharapkan terjadi secara merata hanya berhasil di beberapa tempat yang disebut pusat pertumbuhan. Indikator keberhasilan suatu pembangunan ekonomi didasarkan pada tiga indikator yaitu Indikator moneter, indikator non moneter dan indikator campuran yang selanjutnya dapat digunakan sebagai variabel penelitian (Arsyad, 2010).

Pada skala nasional, Pulau Jawa merupakan pulau dengan peranan distribusi Produk Domestik Regional Bruto terbesar yang kemudian diikuti Pulau Sumatera. Struktur ekonomi Indonesia secara spasial didominasi oleh kelompok provinsi di Pulau Jawa dan Pulau Sumatera. Hal ini didukung dengan data Badan Pusat Statistik yang menunjukkan bahwa Kelompok provinsi di Pulau Jawa memberikan kontribusi terbesar terhadap Produk Domestik Bruto yaitu sebesar 58,49 persen, diikuti oleh Pulau Sumatera sebesar 22,03 persen, dan Pulau Kalimantan 7,85 persen. Dari besarnya kontribusi tersebut Provinsi Jawa Timur juga berperan dalam tingginya tingkat perekonomian. Dikutip dari Berita Resmi Statistik Produk Domestik Regional Bruto

(PDRB) atas dasar harga berlaku di Jawa Timur, mencapai Rp1.855,04 triliun, sedangkan PDRB atas dasar harga konstan mencapai Rp 1.405,24 triliun. Perekonomian Jawa Timur pada tahun 2016 tumbuh sebesar 5,55 persen dari tahun sebelumnya. Jika dibandingkan periode yang sama tahun sebelumnya maka perekonomian tersebut membaik sebesar 5,44 persen (BPS, 2017). Meskipun demikian perekonomian di Jawa Timur belum dapat dikatakan benar-benar baik. Pemerataan ekonomi antar kabupaten/kota masih belum tercapai sehingga timbul masalah ketimpangan ekonomi. Hasil penelitian yang dilakukan Mardiana pada tahun 2012 dengan judul “Kondisi Ketimpangan Ekonomi Antar Kabupaten/Kota dan Implikasinya Terhadap Kebijakan Pembangunan di Provinsi Jawa Timur” menunjukkan tingkat ketimpangan ekonomi antar kabupaten/kota di Provinsi Jawa Timur selama periode analisis tahun 2001-2010 berada pada kesenjangan taraf tinggi. Penelitian selanjutnya oleh Efriza pada tahun 2014 dengan judul “Analisis Kesenjangan Pendapatan antar Kabupaten/Kota di Provinsi Jawa Timur di Era Desentralisasi Fiskal” juga menunjukkan hasil yang hampir sama. Tingkat ketidakmerataan distribusi pendapatan antar kabupaten/kota di Provinsi Jawa Timur begitu tinggi, dimana didapatkan Indeks Williamson yang tinggi. Indeks Williamson merupakan angka digunakan untuk menghitung distribusi pendapatan antar daerah dimana indeks ini telah cukup lazim digunakan dalam mengukur ketimpangan pembangunan antar wilayah. Nilai yang mendekati nol menunjukkan distribusi pendapatan yang merata, sedangkan apabila mendekati 1 maka tingkat ketimpangan ekonomi cukup tinggi.

Upaya mengatasi ketimpangan ekonomi yang terjadi adalah dengan memperlakukan kebijakan-kebijakan khusus pada kabupaten yang memiliki tingkat ekonomi berbeda. Pengelompokan kabupaten atau kota berdasarkan indikator pembangunan ekonomi dapat dilakukan dengan analisis kluster.

Pada umumnya metode klaster mengukur kesamaan antar objek menggunakan ukuran jarak. Namun karena data indikator pembangunan ekonomi mayoritas berupa persentase yang cenderung memiliki banyak data tumpah tindih, maka penggunaan ukuran jarak akan sangat sulit. Oleh karena itu diperlukan metode *cluster* berbeda yang memperhatikan aspek lain misalnya aspek statistik pengamatan. Alternatif metode *cluster* yang dapat digunakan adalah metode *model-based clustering* dimana suatu kelompok pada populasi diidentifikasi berdasar distribusi probabilitas dan keseluruhan populasi dimodelkan sebagai sebuah distribusi *finite mixture* (Banfield & Raftery, 1993). Metode klaster tersebut dianggap lebih baik karena kerangka model probabilitas memberikan struktur yang tetap pada kelompok-kelompok yang terbentuk serta ketepatan pengelompokan yang lebih baik. Hal ini dibuktikan dengan penelitian Pardede (2007) tentang perbandingan metode *K-Means* dan *model-based clustering* yang menunjukkan bahwa persentase kesalahan pengelompokan pada metode *model-based clustering* lebih kecil yaitu 3.33%. Sedangkan persentase kesalahan pengelompokan menggunakan metode *K-means* lebih besar yaitu 10.67%.

Penelitian sebelumnya mengenai pengelompokan kabupaten/kota di Jawa Timur berdasarkan indikator pembangunan ekonomi yang dilakukan Haryono (2016) didapatkan hasil pengelompokan optimum sebanyak 3 klaster. Metode yang digunakan dalam penelitian tersebut adalah *C-Means* dan *Fuzzy C-Means* yang didasarkan pada jarak *Euclidean*. Sedangkan penelitian mengenai *model-based clustering* sebelumnya dilakukan oleh Susilawati (2011) dengan *Gaussian Mixture Models* (GMM) dan Agustini (2017) dengan kriteria *Integrated Completed Likelihood* (ICL) dan *Minimum Message Length* (MML). Dari uraian yang diberikan maka akan dilakukan penelitian untuk mengelompokkan kabupaten/kota di Jawa Timur menjadi kelompok-kelompok tertentu berdasarkan

distribusi probabilitas data indikator pembangunan ekonomi. Hasil penelitian ini diharapkan dapat menjadi salah satu upaya untuk memacu perekonomian di Jawa Timur terutama kabupaten/kota dengan tingkat perekonomian yang rendah. Selain itu pengelompokan kabupaten/kota dapat membantu pemerintah untuk menentukan kebijakan dalam mengurangi ketimpangan ekonomi antar wilayah di Jawa Timur.

1.2 Rumusan Masalah

Rumusan masalah yang akan dibahas pada penelitian ini adalah mengetahui karakteristik dan melakukan pengelompokan kabupaten/kota di Jawa Timur berdasarkan indikator pembangunan ekonomi periode tahun 2011 sampai tahun 2015 dengan analisis *model-based clustering* serta mengetahui perubahan yang terjadi dari tahun ke tahun.

1.3 Tujuan

Berdasarkan rumusan masalah yang telah diuraikan, tujuan penelitian yang ingin dicapai dalam penulisan tugas akhir ini adalah sebagai berikut :

1. Mengetahui karakteristik kabupaten/kota di Jawa Timur berdasarkan indikator pembangunan ekonomi pada tahun 2011 sampai dengan tahun 2015.
2. Melakukan pengelompokan kabupaten/kota di Jawa Timur berdasarkan indikator pembangunan ekonomi pada tahun 2011 sampai dengan tahun 2015 menggunakan metode *model-based clustering*.
3. Mengetahui perubahan pengelompokan kabupaten/kota di Jawa Timur pada tahun 2011 sampai dengan tahun 2015 berdasarkan data indikator pembangunan ekonomi.

1.4 Manfaat

Manfaat yang diperoleh dari penelitian ini secara umum adalah dapat digunakan sebagai referensi untuk sumber

pengetahuan maupun penelitian-penelitian selanjutnya dengan metode maupun studi kasus yang sama. Selanjutnya manfaat khusus dari penelitian ini adalah dapat memberikan gambaran mengenai deskripsi pembangunan ekonomi daerah di Provinsi Jawa Timur pada tahun 2011 sampai dengan tahun 2015 sebagai penilaian perkembangan kinerja pemerintahan pada periode tersebut. Kemudian hasil pengelompokan kabupaten/kota yang terbentuk dapat digunakan sebagai pertimbangan dalam menentukan kebijakan yang tepat secara regional sesuai dengan karakteristik yang khas pada setiap kelompok daerah supaya pemerataan pembangunan di setiap daerah di Jawa Timur dapat terlaksana.

1.5 Batasan Masalah

Pembatasan permasalahan pada kasus ini terdapat pada studi kasus pengamatan indikator pembangunan ekonomi di Jawa Timur yang didasarkan pada tiga indikator utama yaitu indikator moneter, indikator non moneter dan indikator campuran yang didasarkan pada survei sosial ekonomi (Susenas).

BAB II

TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistik deskriptif merupakan metode pengorganisasian yang dapat digunakan dalam penyajian sebuah data secara lebih mudah dan informatif. Proses penyajian data berhubungan dengan hal yang bersifat menguraikan dan memberikan keterangan-keterangan mengenai suatu keadaan atau fenomena. Artinya pada beberapa kasus dapat dikatakan fungsi statistika deskriptif adalah menerangkan keadaan, gejala, atau persoalan yang sedang diteliti (Hasan, 2001).

Penyajian data pada statistika deskriptif diolah dengan grafik, dan penghitungan. Penyajian dengan grafik dapat berupa diagram batang, diagram pencar, histogram, dan lain sebagainya. Sedangkan berdasarkan perhitungan, analisis deskriptif dibagi atas ukuran pemusatan data dan ukuran penyebaran data. Ukuran pemusatan data berupa nilai-nilai yang menunjukkan seberapa besar data memusat pada titik tertentu. Contoh ukuran pemusatan adalah rata-rata (*mean*), nilai tengah (*median*), dan nilai terbanyak (*mode*). Sedangkan ukuran penyebaran merupakan nilai yang menunjukkan seberapa besar persebaran data yang terbentuk dari nilai terkecil dan terbesar. Contoh ukuran penyebaran data adalah varians, simpangan baku, dan nilai jarak (*range*) (Walpole, 2007). Pada penelitian ini statistika deskriptif data dapat disajikan dalam bentuk grafik maupun informasi ukuran pemusatan dan penyebaran dalam bentuk tabel.

2.2 Analisis Cluster

Analisis *cluster* merupakan suatu teknik analisis statistik yang ditujukan untuk menempatkan sekumpulan obyek ke dalam dua atau lebih grup berdasarkan kesamaan-kesamaan obyek atas dasar berbagai karakteristik (Simamora, 2002). Analisis *cluster* berusaha menemukan pengelompokan secara alamiah sesuai

dengan *sense* peneliti. Penempatan objek dilakukan dengan mencari pola dalam kumpulan data agar hasil yang didapatkan optimal. Ciri sebuah *cluster* yang baik adalah *cluster* yang mempunyai homogenitas (kesamaan) yang tinggi antar anggota dalam satu *cluster* (*within cluster*). Sedangkan antar *cluster* satu dengan yang lainnya (*between cluster*) diharapkan memiliki heterogenitas (perbedaan) yang tinggi. Perbedaan analisis *cluster* dengan analisis faktor terletak pada fokus pengelompokan. Analisis *cluster* terfokus pada pengelompokan obyek sedangkan analisis faktor terfokus pada kelompok variabel. Dilihat dari apa yang di*cluster*kan, maka analisis *cluster* dibagi menjadi dua yaitu (Sharma, 1996):

- a. Pengelompokan observasi
- b. Pengelompokan variabel.

Proses *clustering* dipandang sebagai suatu pendekatan yang terbaik untuk menemukan kesamaan pada data sehingga penggunaan analisis *cluster* dalam kehidupan menjadi sangat beragam (Hammouda & Karray, 2000). Beberapa contoh penggunaan *cluster* antara lain masalah segmentasi dalam pemasaran, penelitian mengenai wilayah, permasalahan dalam sistem informasi bahkan penelitian mengenai kesehatan. Sedangkan tujuan utama penggunaan analisis *cluster* adalah (Jain, 2010)

- a. Metode untuk penentuan struktur dalam rangka mendapatkan pola data, membangkitkan hipotesis, mendeteksi adanya penyimpangan, dan mengidentifikasi ciri tertentu yang menonjol.
- b. Metode klasifikasi alamiah untuk menentukan derajat kesamaan antara bentuk-bentuk makhluk hidup (hubungan filogenik).
- c. Metode untuk mendapatkan struktur data yang lebih ringkas dan terorganisasi.

Dasar pengelompokan objek dalam analisis *cluster* adalah suatu ukuran yang menyatakan korespondensi antar objek. Ukuran tersebut dinyatakan dalam ukuran kesamaan yaitu ukuran

korelasi, ukuran jarak, dan ukuran asosiasi. Ukuran korelasi dapat diterapkan pada data dengan skala metrik, tetapi jarang digunakan karena titik beratnya pada nilai suatu pola tertentu. Ukuran jarak merupakan ukuran ketidakmiripan, dimana jarak yang besar menunjukkan sedikit kesamaan, sedangkan jarak yang pendek/kecil menunjukkan bahwa suatu objek makin mirip dengan objek lain. Beberapa tipe ukuran jarak antara lain jarak *euclidean*, jarak *city-block (manhattan)*, dan jarak *mahalanobis*. Kemudian ukuran asosiasi dipakai untuk mengukur data berskala nonmetrik (nominal atau ordinal).

Secara umum analisis *cluster* dibedakan menjadi dua yaitu metode hierarki (*hierarchical clustering*) dan metode non-hierarki (*non-hierarchical clustering*) (Johnson & Wichern, 2007). *Cluster* yang dibentuk dengan metode hirarki dilakukan tanpa menentukan jumlah kelompok terlebih dahulu. Jumlah kelompok beserta pengelompokannya akan terbentuk dari pendekatan-pendekatan yang dilakukan. Jarak yang digunakan dalam analisis ini adalah jarak *Eucliden* sebagai berikut.

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2} \quad (2.1)$$

Beberapa metode penggabungan yang bisa digunakan dalam analisis *cluster* hirarki antara lain metode pautan tunggal (*single linkage*), pautan lengkap (*complete linkage*), pautan rata-rata (*average linkage*), metode *ward (ward's method)* dan metode *centroid (centroid method)*.

Metode non hirarki dipakai jika banyaknya *cluster* sudah diketahui dan biasanya metode ini dipakai untuk mengelompokkan data yang berukuran besar (Hair et al, 2010). Perhatian utama dalam metode non hirarki adalah bagaimana memilih awalan atau inisial *cluster* yang berpengaruh besar terhadap hasil akhir analisis *cluster*. Adapun yang termasuk dalam metode ini adalah metode *K-means*, *K-Modes*, *Density Estimation* dan *Mixtures of Distribution*. Metode *K-means* hanya dapat digunakan dalam data numerik sehingga untuk data yang sifatnya kategorik dikembangkan sebuah metode yang disebut *K-*

Modes. *K-modes* memiliki dasar pemikiran yang sama dengan *K-means* namun menggunakan modus data sebagai pusat centroid. (Huang & Michael, 1999). Sedangkan *Density Estimation* melakukan pengelompokan dengan mencari wilayah yang memiliki kepadatan (*density*) yang tinggi. Estimasi *density* dilakukan menggunakan fungsi kernel dan tidak memerlukan asumsi khusus. Kemudian untuk *Mixtures of Distribution* mendasari identifikasi kelompok melalui distribusi probabilitas. Keseluruhan populasi dimodelkan sebagai sebuah *mixture* dari distribusi. Sehingga dapat dikatakan pengelompokan pada metode ini adalah pengelompokan berdasarkan model dari distribusi probabilitas.

2.3 Distribusi t Multivariat

Dalam satu percobaan selalu terdapat elemen ruang sampel dan himpunan bagian dari ruang sampel yang disebut kejadian atau *events*. Seberapa sering *events* tersebut terjadi didalam ruang sampel tersebut ditandai sebagai suatu peluang atau probabilitas dimana nilainya berada diantara 0 dan 1 (Dekking et al, 2005). Distribusi peluang sangat berguna dalam analisis terjadinya suatu peristiwa atau kejadian. Dengan mendefinisikan variabel acak sebagai fungsi tertentu pada setiap distribusi maka kejadian pada ruang sampel dapat dipetakan ke dalam himpunan bilangan real (Casella & Berger, 2002). Secara umum distribusi probabilitas terbagi menjadi distribusi peluang diskrit dan distribusi peluang kontinyu. Distribusi peluang diskrit adalah distribusi peluang yang terjadi ketika ruang contoh mengandung titik contoh yang banyaknya sama dengan bilangan cacah. Sedangkan distribusi peluang kontinyu adalah distribusi peluang yang terjadi ketika ruang contoh mengandung titik contoh yang jumlahnya tak berhingga sehingga peluang untuk mendapat satu titik adalah nol (Walpole, 2007).

Distribusi t multivariat merupakan alternatif distribusi pengganti ketika pada data terdapat banyak outlier, maka distribusi data menjadi lebih landai sehingga menjadi tidak normal multivariat. Distribusi t multivariat adalah perluasan dari

distribusi t univariat. Jika vektor variabel acak $\mathbf{x} = [x_1 x_2 \dots x_p]^T$ berdistribusi t multivariat dengan derajat bebas v , maka vektor rata-rata $\boldsymbol{\mu} = [\mu_1 \mu_2 \dots \mu_p]^T$ dan matriks varians kovarian Σ memiliki fungsi kepadatan peluang sebagai berikut :

$$f(\mathbf{x}) = \frac{\Gamma\left(\frac{v+p}{2}\right)}{(\pi v)^{\frac{p}{2}} \Gamma\left(\frac{v}{2}\right) |\Sigma|^{\frac{1}{2}}} \left(1 + \frac{(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})}{v}\right)^{-\left(\frac{v+p}{2}\right)}; v > 2 \quad (2.2)$$

Dimana v disebut juga parameter bentuk karenadengan memberikan variasi nilai pada v maka fungsi pada persamaan 2.4 dapat diturunkan atau dinaikkan. Distribusi t untuk data multivariat yang diakui memiliki kemampuan mengatasi *outlier* lebih mudah dibandingkan distribusi normal multivariat sehingga distribusi ini menjadi sering digunakan dalam *model-based clustering* (Andrews et al, 2011).

2.4 Deteksi *Outlier* Multivariat

Data tidak bersitribusi normal multivariat atau matriks varians kovariansnya tidak homogen dapat disebabkan adanya observasi yang mempunyai pola berbeda dengan sebagian besar pola data. Observasi tersebut disebut data pencilan (*outlier*). Dalam kasus pengelompokan dengan metode *model-based clustering*, *outlier* dapat menjadikan hasil pengelompokan kurang tepat dan penaksir parameter menjadi bias. Oleh sebab itu, sangat penting memeriksa keberadaan *outlier*. Metode yang digunakan untuk memeriksa keberadaan *outlier* multivariat adalah perhitungan Jarak Mahalanobis yang didefinisikan sebagai berikut :

$$MD_i = [(\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})]^{1/2}, i = 1, 2, \dots, n \quad (2.3)$$

dimana $\bar{\mathbf{x}}$ adalah vektor *mean* sampel dan \mathbf{S} adalah matriks varians kovarians sampel. Hasil deteksi suatu observasi dikatakan data *outlier* jika nilai akar kuadrat jarak Mahalanobis lebih besar dari $\chi_{p;1-\alpha/2}^2$ dimana p adalah derajat bebas (Rousseeuw & Van Zomeren, 1990).

2.5 Model Based-Clustering

Model-based clustering adalah metode pengelompokan yang mengoptimalkan kemiripan antara objek berdasarkan pada pendekatan model probabilistik. Metode ini merupakan pengembangan metode klasik yang umumnya menggunakan jarak *euclidian* dalam pengelompokan. Pendekatan tersebut dapat memodelkan data yang dimiliki dengan menerapkan pengaturan karakteristik yang berbeda-beda dan menentukan jumlah *cluster* yang sesuai dengan data seiring proses pemodelan karakteristik dari masing-masing kelompok tersebut. Model ini pertama kali digunakan oleh Banfield & Raftery (1993) untuk pengelompokan objek dalam populasi. *Model-based clustering* mengasumsikan bahwa data dibangkitkan oleh campuran dari sebaran peluang dengan masing-masing komponen mewakili *cluster* berbeda sehingga objek didalam suatu *data set* dapat dikelompokkan menjadi kelompok data yang sebelumnya tidak terdefinisi. Kelebihan metode *cluster* ini adalah struktur pada kelompok-kelompok lebih stabil serta ketepatan pengelompokan yang lebih baik karena kelompok dibentuk secara alami dari distribusi data tanpa ditentukan jumlahnya. Asumsi yang digunakan pada *model-based clustering* adalah dalam suatu populasi dapat diambil subpopulasi yang mempunyai distribusi peluang tertentu dan masing-masing subpopulasi mempunyai parameter yang berbeda. Kerangka *model-based clustering* dikembangkan berdasarkan dekomposisi *eigenvalue* dari matriks varians kovarians (Σ_g) sebagai berikut :

$$\Sigma_g = \lambda_g D_g A_g D_g^T \quad (2.4)$$

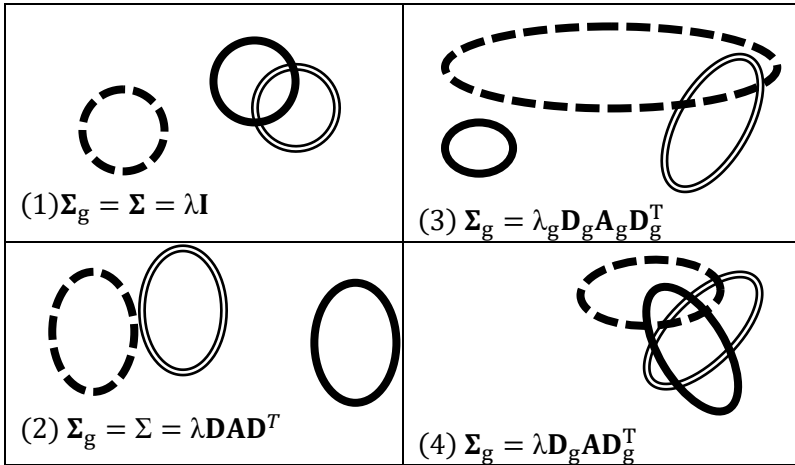
dimana :

λ_g adalah nilai skalar yang menunjukkan volume elips

D_g adalah matriks ortogonal *eigenvector* yang merupakan orientasi komponen utama Σ_g

A_g adalah matriks diagonal dengan elemen-elemen yang proporsional pada *eigenvalue* Σ_g dan menunjukkan kontur dari fungsi kepadatannya

Dengan bantuan *software R* pada *package teigen* dapat dilakukan pengelompokan menggunakan *model-based clustering t* multivariat. Jumlah model yang diidentifikasi dapat terbentuk adalah 28 model dengan jumlah kelompok maksimal 9 kelompok. Model terbentuk akan dinotasikan dengan huruf-huruf untuk mewakili kerangka model yaitu label C (*constrained*), I (matriks identitas) dan U (*unconstrained*) yang secara lengkap berada pada Lampiran 11. Selanjutnya macam-macam hasil orientasi, volume dan bentuk dari distribusi yang diestimasi dari data akan diilustrasikan pada Gambar 2.1



Gambar 2.1 Ilustrasi Bentuk *Cluster* Berdasarkan Matriks Varians Kovarians

Penjelasan ilustrasi pada Gambar 2.1 adalah sebagai berikut.

1. $\Sigma_g = \Sigma = \lambda \mathbf{I}$ menghasilkan semua *cluster* berbentuk bola dengan volume sama besar antar *cluster*
2. $\Sigma_g = \Sigma = \lambda \mathbf{D} \mathbf{A} \mathbf{D}^T$ menghasilkan *cluster* yang mempunyai bentuk, volume dan orientasi yang sama
3. $\Sigma_g = \lambda_g \mathbf{D}_g \mathbf{A}_g \mathbf{D}_g^T$ menghasilkan *cluster* yang berbeda baik pada bentuk, volume maupun orientasinya.
4. $\Sigma_g = \lambda \mathbf{D}_g \mathbf{A} \mathbf{D}_g^T$ menghasilkan *cluster* yang hanya hanya berbeda pada orientasinya saja (Agustini, 2017).

2.6 Model Finite Mixture

Misalkan vektor variabel acak x dengan dimensi p berasal dari distribusi *finite mixture* dengan fungsi kepadatan peluang :

$$f(\mathbf{x}|\boldsymbol{\vartheta}) = \sum_{g=1}^G \pi_g f_g(\mathbf{x}|\boldsymbol{\theta}_g) \quad (2.5)$$

dimana $\boldsymbol{\vartheta} = (\pi_1, \pi_2, \dots, \pi_G, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_G)$ adalah vektor parameter. $f_g(x|\boldsymbol{\theta}_g)$ disebut fungsi kepadatan peluang \mathbf{x} dengan parameter kelompok $\boldsymbol{\theta}_G$. Sedangkan G adalah jumlah kelompok dan π_g adalah bobot atau proporsi campuran (*mixing proportion*) dari kelompok ke- g dengan batasan :

$$0 \leq \pi_j \leq 1, \quad g = 1, 2, \dots, G \text{ dan } \sum_{g=1}^G \pi_g = 1$$

Tabel 2.1 Struktur Data Kelompok ke- g pada Model *Finite Mixture*

Objek	Variabel					
	1	2	...	j	...	p
1	$(x_{11})_g$	$(x_{21})_g$...	$(x_{j1})_g$...	$(x_{p1})_g$
2	$(x_{12})_g$	$(x_{22})_g$...	$(x_{j2})_g$...	$(x_{p2})_g$
\vdots	\vdots	\vdots		\vdots		\vdots
i	$(x_{1i})_g$	$(x_{2i})_g$...	$(x_{ji})_g$...	$(x_{pi})_g$
\vdots	\vdots	\vdots		\vdots		\vdots
n_g	$(x_{1n_g})_g$	$(x_{2n_g})_g$...	$(x_{jn_g})_g$...	$(x_{pn_g})_g$

dimana

$(x_{ji})_g$: Nilai variabel ke- j pengamatan ke- i kelompok ke- g

$(\mathbf{x}_i)_g$: Vektor nilai data objek pengamatan ke- i kelompok ke- g

g : Indeks kelompok, $g = 1, 2, \dots, G$

j : Indeks variabel, $j = 1, 2, \dots, p$

i : Indeks objek pengamatan, $i = 1, 2, \dots, n_g$

Model *finite mixture* menggunakan dua proses utama yaitu penaksiran parameter dan pemilihan model terbaik. Penaksiran parameter *model-based clustering* dilakukan dengan algoritma *Expectation Maximization* (EM) sedangkan pemilihan model terbaik dilakukan berdasarkan kriteria *Integrated Completed Likelihood* (ICL) (Bienarcki et al, 2000).

Pada kasus untuk data mengandung *outlier* model *finite mixture t* multivariat dianggap sebagai sebuah pendekatan yang lebih *robust*. Dengan menganggap setiap kelompok pada finite mixture berdistribusi *t* multivariat maka efek *outlier* dapat diatasi. Distribusi *t* memiliki ekor yang lebih panjang dibanding distribusi normal sehingga data pengamatan yang tidak biasa dari sebuah kelompok akan mendapat bobot yang lebih rendah dalam penghitungan nilai penaksir parameternya (McLachlan & Peel, 2000).

Jika sampel acak $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ berasal dari sebuah model *finite mixture t* multivariat dengan setiap kelompok ke- g pada model diasumsikan berdistribusi *t* multivariat maka fungsi kepadatan peluang yang terbentuk adalah sebagai berikut

$$f(\mathbf{x}_i; \pi_g; \boldsymbol{\mu}_g; \boldsymbol{\Sigma}_g; v_g) = \sum_{g=1}^G \pi_g f_g(\mathbf{x}_i; \boldsymbol{\mu}_g; \boldsymbol{\Sigma}_g; v_g) \quad (2.6)$$

dimana $g = 1, 2, \dots, G; i = 1, 2, \dots, n; \sum_{g=1}^G \pi_g = 1$ dan

$$f_g(\mathbf{x}_i; \boldsymbol{\mu}_g; \boldsymbol{\Sigma}_g; v_g) = \frac{\Gamma\left(\frac{v_g+p}{2}\right) |\boldsymbol{\Sigma}_g|^{-\frac{1}{2}}}{(\pi v_g)^{\frac{p}{2}} \Gamma\left(\frac{v_g}{2}\right)} \left(1 + \frac{\delta(\mathbf{x}_i; \boldsymbol{\mu}_g; \boldsymbol{\Sigma}_g)}{v_g}\right)^{-\left(\frac{v_g+p}{2}\right)} \quad (2.7)$$

dengan

$$\delta(\mathbf{x}_i; \boldsymbol{\mu}_g; \boldsymbol{\Sigma}_g) = (\mathbf{x}_i - \boldsymbol{\mu}_g)^T \boldsymbol{\Sigma}_g^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_g) \quad (2.8)$$

adalah jarak Mahalanobis kuadrat antara \mathbf{x}_i dan $\boldsymbol{\mu}_g$ ($\boldsymbol{\Sigma}$ sebagai matriks kovarian). Ketika derajat bebas kelompok ke- g mendekati tak hingga ($v_g \rightarrow \infty$) maka pendekatan dengan distribusi *t* ini akan mengarah pada distribusi normal. Sehingga parameter v_g dianggap sebagai parameter pengontrol *robust* (*robustness tuning*) (McLachlan & Peel, 2000).

2.7 Algoritma Expectation Maximization

Algoritma *Expectation Maximization* merupakan algoritma yang digunakan untuk mempermudah mendapatkan penaksir parameter pada *model-based clustering* (Dempster et al, 1997). Aplikasi yang paling tepat dari algoritma tersebut adalah untuk pendekatan umum *Maximum Likelihood* pada data yang tidak lengkap. Namun data observasi dapat dianggap sebagai data tidak

lengkap ketika label kelompok belum diketahui, sehingga diperlakukan sebagai variabel laten. Jika label ini diketahui, maka akan didapatkan estimasi parameter di setiap distribusi komponen dengan membagi observasi ke dalam kelompok tertentu. Dalam Algoritma EM, data terdiri dari n pengamatan variabel ganda yang diperoleh dari $(\mathbf{y}_i^T, \mathbf{z}_i^T)$ dimana y_i merupakan variabel yang teramati dan z_i merupakan variabel yang tidak teramati. Variabel tersebut yang akan menempatkan objek masuk ke dalam kelompok tertentu. Jika x_i saling bebas dan terdistribusi identik menurut sebaran peluang f dengan parameter θ , maka *likelihood* data lengkap adalah

$$L_c(\mathbf{x}_i|\boldsymbol{\theta}) = \prod_{i=1}^n f(\mathbf{x}_i|\boldsymbol{\theta}) \quad (2.9)$$

Sedangkan untuk *likelihood* data tak lengkap adalah

$$L_o(\mathbf{y}|\boldsymbol{\theta}) = \int L_c(\mathbf{x}|\boldsymbol{\theta}) dz \quad (2.10)$$

Algoritma EM menggunakan iterasi 2 langkah sampai mencapai hasil yang konvergen, yaitu langkah Ekspektasi (*E-step*) dan langkah Maksimisasi (*M-step*). Pada *E-step* nilai harapan bersyarat dari fungsi *loglikelihood* data lengkap akan dihitung menggunakan penduga parameternya berdasarkan rumus berikut:

$$\ln L_c(\vartheta|y_i) = \sum_{i=1}^n \sum_{g=1}^G z_{ig} \left[\ln(\pi_g) + \ln(f_g(x_i|\theta_g)) \right] \quad (2.11)$$

Berdasarkan ekspektasi dari fungsi 2.11 didapatkan penaksir parameter

$$\hat{z}_{ig} = \frac{\hat{\pi}_g f_g(x_i|\theta_g)}{\sum_{g=1}^G \hat{\pi}_g f_g(x_i|\theta_g)} \quad (2.12)$$

Selanjutnya pada *Maximization Step* (*M-Step*) akan dilakukan perhitungan parameter yang memaksimalkan nilai harapan dari fungsi *log likelihood* terhadap $\hat{\pi}_g$, $\hat{\mu}_g$ dan $\hat{\Sigma}_g$ dengan \hat{z}_{ig} yang diperoleh pada *E-Step*.

$$\hat{z}_g = \frac{n_g}{n} \text{ dan } \hat{\mu}_g = \frac{\sum_{i=1}^n z_{ig} y_i}{n_g} \quad (2.13)$$

Algoritma EM dilakukan terus-menerus sampai mendapatkan hasil iterasi yang konvergen sehingga anggota *cluster* dapat dikelompokkan menggunakan metode klasifikasi *Maximum a Posteriori* (MAP) dimana dengan syarat sebagai berikut:

$$\text{MAP}\{\hat{z}_{ig}\} = \begin{cases} 1 & \text{jika } \max\{\hat{z}_{ig}\} \in k - g \\ 0 & \text{lainnya} \end{cases} \quad (2.14)$$

2.8 Integrated Completed Likelihood

Integrated Completed Likelihood merupakan salah satu kriteria yang dapat digunakan dalam memilih model terbaik dari *model-based clustering* (Bienarcki et al, 2000). Pemilihan model terbaik berkaitan erat dengan estimasi jumlah kelompok yang diperoleh. Hubungan antara dua hal tersebut terbalik karena ketika pemilihan model terbaik dilakukan dengan model sederhana akan didapatkan banyak kelompok yang menggambarkan struktur data. Sebaliknya jika model yang dipilih semakin kompleks maka akan semakin sedikit kelompok yang memenuhi kriteria yang ditentukan (Fraley & Raftery, 2002). Dalam menggunakan kriteria *Integrated Completed Likelihood* fungsi likelihood data lengkap $y_i = (x_i, z_i)$ akan dimaksimumkan, sehingga rumus ICL yang didapatkan adalah sebagai berikut :

$$ICL_g = \ln f(y_i) - \frac{p}{2} \ln(n) \quad (2.15)$$

dimana $f(y_i) = f(x_i, y_i)$ adalah fungsi kepadatan peluang bersama data lengkap, p adalah banyaknya parameter, dan n adalah banyaknya observasi.

Pada penelitian yang dilakukan oleh Bienarcki et al (2000) dan Baudry et al (2013) didapatkan kesimpulan bahwa ICL mampu memberikan perkiraan jumlah komponen *mixture* yang stabil dan reliabel baik pada data realita maupun data simulasi, dilihat dari sudut pandang pengelompokan. Akan tetapi, ICL dapat mengabaikan jumlah komponen untuk data simulasi yang timbul dari komponen *mixture* yang tidak terpisah dengan baik (*overlap*). Model terbaik yang dipilih pada kriteria ICL adalah model dengan nilai ICL paling besar.

2.9 Uji Manova

Dalam suatu analisis seringkali lebih dari satu populasi ingin dibandingkan agar dapat diketahui perbedaan antar perlakuan

Untuk mengetahui apakah vektor rata-rata populasi sama maka dilakukan uji *multivariate analysis of variance* (manova) (Johnson & Wichern, 2007). Uji ini juga dapat digunakan untuk mengetahui komponen rata-rata mana yang berbeda secara nyata jika hasil secara serentak menunjukkan perbedaan. Kemudian sampel acak dikumpulkan dari populasi yang disusun sebagai berikut :

Populasi 1 : $x_{11}, x_{12}, \dots, x_{1n1}$

Populasi 2 : $x_{21}, x_{22}, \dots, x_{2n2}$

⋮

Populasi G : $x_{G1}, x_{G2}, \dots, x_{Gng}$

Model manova untuk membandingkan G vektor rata-rata populasi adalah

$$\mathbf{x}_{gi} = \boldsymbol{\mu} + \boldsymbol{\tau}_g + \boldsymbol{\varepsilon}_{gi} \quad (2.16)$$

dengan $i = 1, 2, \dots, n$ dan $g = 1, 2, \dots, G$, $\boldsymbol{\mu}$ adalah vektor rata-rata keseluruhan dan $\boldsymbol{\tau}_g$ menyatakan pengaruh perlakuan ke- g . Adapun hipotesis yang digunakan dalam pengujian manova adalah sebagai berikut.

$H_0: \tau_1 = \tau_2 = \dots = \tau_G = 0$ (tidak ada perbedaan antar kelompok)

H_1 : minimal satu $\tau_i \neq \tau_j$ untuk $i \neq j$ (ada perbedaan antar kelompok)

Statistik Uji menggunakan kriteria *Wald Type* adalah sebagai berikut (Bathke et al, 2016)

$$Q_N(T) = N\bar{\mathbf{X}}^T T (T\hat{\mathbf{V}}_N T) + T\bar{\mathbf{X}} \quad (2.17)$$

dengan

$$\bar{\mathbf{X}} = (\bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_d)^T, \bar{X}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} X_{i,k} \text{ dan } \hat{\mathbf{V}}_N = \text{diag} \left(\frac{N}{n_i} \right) \hat{\Sigma}_i, \hat{\Sigma}_i = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (X_{ik} - \bar{X}_i)(X_{ik} - \bar{X}_i)^T, 1 \leq i \leq d$$

Keputusan tolak H_0 jika $p - \text{value} < \alpha$ (0,05) yang berarti ada perbedaan antar kelompok yang terbentuk.

2.10 Pembangunan Ekonomi

Pembangunan ekonomi dapat didefinisikan sebagai usaha untuk meningkatkan pendapatan perkapita dengan jalan mengolah kekuatan ekonomi potensial menjadi ekonomi riil melalui

penggunaan teknologi, penambahan pengetahuan, peningkatan ketrampilan, penambahan kemampuan berorganisasi dan manajemen (Sukirno, 1996). Pembangunan dan pertumbuhan ekonomi bukanlah sesuatu yang sama namun keduanya memiliki keterkaitan. Pertumbuhan ekonomi lebih menekankan pada proses peningkatan produksi barang dan jasa dalam kegiatan ekonomi masyarakat. Sedangkan pembangunan ekonomi mengandung pengertian yang lebih luas dan mencakup perubahan pada tata susunan ekonomi masyarakat secara menyeluruh (Sumitro, 1987). Pada umumnya pembangunan selalu disertai dengan pertumbuhan, tetapi pertumbuhan belum tentu disertai dengan pembangunan. Pada tingkat permulaan, pembangunan ekonomi dibarengi pula dengan pertumbuhan dan sebaliknya (Irawan & Suparmoko, 1988). Sehingga secara sederhana dapat dikatakan bahwa pertumbuhan tersebut merupakan salah satu indikator keberhasilan pembangunan ekonomi. Selanjutnya sasaran pembangunan ekonomi menurut Todaro (2010) adalah :

- a. Meningkatkan persediaan dan memperluas pembagian/pemerataan bahan pokok yang dibutuhkan untuk bisa hidup, seperti kelayakan tempat tinggal, kesehatan dan lingkungan.
- b. Mengangkat taraf hidup, menambah dan mempertinggi pendapat dan penyediaan lapangan kerja, pendidikan yang lebih baik, yang semata-mata bukan hanya untuk memenuhi kebutuhan materi, tetapi untuk meningkatkan kesejahteraan baik individu maupun nasional.
- c. Memperluas jangkauan pilihan ekonomi dan sosial bagi semua individu dan nasional dengan cara membebaskan mereka dari sikap budak dan ketergantungan, tidak hanya hubungan dengan orang lain dan negara lain, tetapi juga dari sumber-sumber kebodohan dan penderitaan.

Pada lingkup yang lebih kecil, pembangunan ekonomi pada daerah atau wilayah tertentu disebut pembangunan ekonomi daerah. Pembangunan ekonomi daerah melibatkan pemerintah daerah dan masyarakat dalam mengelola sumber daya yang ada dan membentuk suatu pola kemitraan antara pemerintah daerah

dengan sektor swasta untuk mendorong pertumbuhan ekonomi di wilayah tertentu (Arsyad, 1999). Namun keadaan sosial ekonomi yang berbeda di setiap daerah justru berdampak kurang baik. Perbedaan tingkat pembangunan antar daerah, mengakibatkan perbedaan tingkat kesejahteraan daerah sehingga terdapat pemusatan ekspansi ekonomi. Memusatnya ekspansi ekonomi dapat disebabkan beberapa faktor, misalnya konsisi dan situasi alamiah yang ada, letak geografis, maupun kondisi pemerintahan. Adanya hal tersebut merugikan daerah lain karena beberapa faktor ekonomi seperti tenaga kerja, modal dan proses perdagangan, akan berpindah pada daerah tertentu saja (Jhingan, 1993). Akibatnya daerah-daerah yang terbelakang akan semakin tertinggal dan pada akhirnya dapat mengganggu kestabilan ekonomi negara secara keseluruhan.

Untuk memberikan suatu gambaran mengenai hasil pembangunan ekonomi maka diperlukan indikator-indikator yang dapat digunakan sebagai acuan. Indikator-indikator pembangunan ekonomi terbagi menjadi indikator moneter, indikator non moneter dan campuran (Arsyad, 2010). Secara umum indikator yang dalam pembangunan ekonomi adalah sebagai berikut

a. Indikator Moneter

Indikator moneter berkaitan dengan tingkat pendapatan masyarakat. Indikator tersebut mencakup pendapatan per kapita dan indikator kesejahteraan ekonomi bersih. Pengukuran indikator dapat dilakukan melalui Produk Domestik Regional Bruto (PDRB) yang merupakan jumlah keseluruhan nilai tambah barang dan jasa yang dihasilkan dari semua kegiatan perekonomian di suatu wilayah dan dalam periode tahun tertentu. Perhitungan PDRB biasanya dilakukan dalam waktu satu tahun. Dalam perhitungan PDRB terdapat empat pendekatan yaitu pendekatan produksi, Pendekatan pendapatan, Pendekatan pengeluarandan Metode alokasi.

Dalam perhitungan PDRB dikenal istilah PDRB perkapita yang merupakan hasil dari pembagian PDRB dengan jumlah penduduk pertengahan tahun. Berikut adalah perumusannya

$$PDRB \text{ perkapita} = \frac{PDRB \text{ tahun } x}{Jumlah \text{ penduduk tahun } x} \quad (2.18)$$

PDRB atas dasar harga konstan menunjukkan nilai tambah barang dan jasa tersebut yang dihitung menggunakan harga yang berlaku pada satu tahun tertentu sebagai dasar, sehingga harga konstan digunakan untuk mengetahui pertumbuhan ekonomi dari tahun ke tahun. Dalam perhitungan PDRB atas dasar harga berlaku, semua agregat pendapatan dinilai atas dasar harga berlaku pada tahun berjalan. Untuk mendapatkan Nilai Tambah Bruto atas dasar harga Berlaku (NTBB) guna perhitungan PDRB atas dasar harga berlaku digunakan rumus sebagai berikut

$$NTBB_{i,s} = \frac{NTBK_{i,s-1} \times IH_{i,s}}{100} \quad (2.19)$$

Dimana :

$NTBB_{i,s}$: Nilai tambah bruto atas dasar harga berlaku sektor ke-i semester s

$NTBK_{i,s-1}$: nilai tambah bruto atas dasar harga konstan sektor ke-i semester s-1

$IH_{i,s}$: Indeks harga sektor ke-i semester s

b. Indikator Non-Moneter

Indikator non-moneter berkaitan dengan kehidupan masyarakat yang meliputi indikator sosial dan Indeks Pembangunan Manusia (IPM). Indeks Pembangunan Manusia adalah indeks untuk yang digunakan untuk mengukur perkembangan pembangunan manusia dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk). Selain itu indeks tersebut dapat digunakan untuk menentukan peringkat atau level pembangunan suatu wilayah/negara. Dasar pengukuran Indeks Pembangunan Manusia didasarkan pada aspek kesehatan, pendidikan dan kemampuan secara ekonomi sehingga rumus perhitungan IPM adalah sebagai berikut

$$IPM = \sqrt[3]{I_{kesehatan} \times I_{pendidikan} \times I_{pengeluaran}} \quad (2.20)$$

c. Indikator Campuran

Indikator campuran merupakan indikator yang terdiri atas indikator sosial dan ekonomi. Indikator ini adalah hasil pengembangan indikator kesejahteraan rakyat oleh Badan Pusat

Statistik pada Survei Sosial Ekonomi (SUSENAS). Indikator Susenas Inti ini meliputi aspek-aspek sebagai berikut

a. Pendidikan

Pendidikan dapat digunakan dalam mengukur pembangunan ekonomi suatu wilayah. Tingkat pendidikan diukur berdasarkan Angka Partisipasi Sekolah (APS), yaitu angka yang menyatakan daya tampung sarana pendidikan dalam memfasilitasi penduduk berdasarkan usia tertentu untuk menikmati pendidikan. APS menunjukkan seberapa besar penduduk di suatu wilayah dapat merasakan fasilitas pendidikan. Perhitungan Angka Partisipasi Sekolah pada usia tertentu dapat dilakukan dengan rumus berikut

$$APS = \frac{\text{jumlah penduduk yang sekolah}}{\text{jumlah total penduduk}} \times 100\% \quad (2.21)$$

Semakin tinggi Angka Partisipasi Sekolah semakin besar jumlah penduduk yang berkesempatan mengenyam pendidikan. Namun demikian meningkatnya APS tidak selalu dapat diartikan bahwa pemerataan kesempatan masyarakat untuk mengenyam pendidikan sudah baik.

b. Kesehatan

Kesehatan merupakan hak setiap orang agar dapat merasa nyaman dalam kehidupan bermasyarakat. Indikator tingkat kesehatan dapat dilihat dari rata-rata hari sakit maupun ketersediaannya fasilitas kesehatan. Untuk perhitungan tingkat keluhan kesehatan digunakan rumus berikut

$$TKK = \frac{\text{jumlah penduduk yang keluhan kesehatan}}{\text{jumlah total penduduk}} \times 100\% \quad (2.22)$$

c. Perumahan

Salah satu kebutuhan utama yang harus terpenuhi oleh setiap penduduk adalah rumah. Kelayakan hunian merupakan salah satu indikator keberhasilan pembangunan ekonomi. Tingkat kelayakan dapat dilihat dari sumber air bersih, Sumber penerangan listrik, sanitasi, dan mutu rumah tinggal. Rumus perhitungan variabel persentase ketersediaan listrik adalah

$$PKL = \frac{\text{jumlah penduduk memiliki sumber listrik}}{\text{jumlah total penduduk}} \times 100\% \quad (2.23)$$

d. Angkatan Kerja

Tingkat kebaikan angkatan kerja dapat dilihat dari Tingkat Pengangguran Terbuka dan Tingkat Partisipasi Angkatan Kerja. Tingkat Pengangguran Terbuka (TPT) adalah indikasi tentang penduduk usia kerja yang termasuk dalam kelompok pengangguran. Tingkat pengangguran terbuka diukur sebagai persentase jumlah penganggur/pencari kerja terhadap jumlah angkatan kerja, yang dapat dirumuskan sebagai berikut

$$TPT = \frac{\text{jumlah pengangguran}}{\text{jumlah angkatan kerja}} \times 100\% \quad (2.24)$$

indikator pengangguran terbuka berguna sebagai acuan pemerintah bagi pembukaan lapangan kerja baru. Perkembangan nilai tersebut menunjukkan tingkat keberhasilan program ketenagakerjaan dari tahun ke tahun.

Tingkat Partisipasi Angkatan Kerja (TPAK) adalah perbandingan antara angkatan kerja dengan jumlah seluruh penduduk usia kerja. TPAK mengukur besarnya partisipasi angkatan kerja dalam dunia kerja. Rumus perhitungan TPAK adalah

$$TPAK = \frac{\text{jumlah angkatan kerja}}{\text{jumlah penduduk usia kerja}} \times 100\% \quad (2.25)$$

Angka persentase yang rendah menunjukkan kecilnya kesempatan kerja yang tersedia bagi penduduk usia kerja, sedangkan angka yang tinggi menunjukkan besarnya kesempatan kerja yang tersedia .

e. Keluarga Berencana dan Fertilitas

Indikator dalam bagian ini antara lain penggunaan asi, tingkat imunisasi, kehadiran tenaga kesehatan pada kelahiran dan penggunaan alat kontrasepsi.

f. Ekonomi

Indikator yang diamati dalam bagian survei sosial ekonomi adalah tingkat pendapatan dan konsumsi per kapita. Selain itu kemiskinan yang didefinisikan sebagai permasalahan ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan juga dapat dimasukkan

dalam bagian ini. Rumus perhitungan presentase kemiskinan adalah sebagai berikut

$$PK = \frac{\text{jumlah penduduk miskin}}{\text{jumlah total penduduk}} \times 100\% \quad (2.26)$$

g. Kriminalitas

Tingkat kriminalitas dihubungkan dengan kualitas pembangunan ekonomi. Ketika pembangunan ekonomi sudah berhasil maka tingkat kriminalitas akan rendah. Indikator yang digunakan dalam bagian ini adalah persentase tindak pidana terselesaikan. Peristiwa yang dilaporkan ialah setiap peristiwa yang diterima kepolisian dari laporan masyarakat, atau peristiwa dimana pelakunya tertangkap tangan oleh kepolisian. Untuk itu persentase penyelesaian peristiwa kejahatan dapat dirumuskan sebagai berikut.

$$PPK = \frac{\text{jumlah kejahatan terselesaikan}}{\text{jumlah kejahatan dilaporkan}} \times 100\% \quad (2.27)$$

h. Perjalanan Wisata

Sektor pariwisata dapat digunakan sebagai indikator pembangunan ekonomi dimana semakin besar frekuensi kunjungan wisata pada tiap daerah menunjukkan tingkat pembangunan yang semakin baik.

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Penelitian ini menggunakan data sekunder yang diperoleh dari Badan Pusat Statistika Provinsi Jawa Timur mengenai data indikator pembangunan ekonomi. Data tersebut merupakan bagian dari publikasi data Survei Sosial Ekonomi Nasional (Susenas) tahun 2011 sampai dengan tahun 2015. Survei Sosial Ekonomi Nasional merupakan kegiatan survei mengenai data dibidang kependudukan, kesehatan, pendidikan, Keluarga Berencana, perumahan, serta konsumsi dan pengeluaran. Data tersebut merupakan data dengan unit pengamatan yang diambil pada tingkat Kabupaten/ Kota di Provinsi Jawa Timur sebanyak 38 Kabupaten/ Kota.

3.2 Variabel Penelitian

Adapun variabel penelitian yang digunakan berdasarkan referensi dari penelitian sebelumnya ditampilkan pada Tabel 3.1 berikut ini

Tabel 3.1 Variabel Penelitian

Nama	Keterangan	Skala	Satuan
X ₁	Produk Domestik Regional Bruto perkapita	Rasio	Trilyun Rupiah
X ₂	Produk Domestik Bruto atas harga berlaku	Rasio	Trilyun Rupiah
X ₃	Indek Pembangunan Manusia	Rasio	Persen
X ₄	Angka Partisipasi Sekolah berusia 7 sampai 12 tahun	Rasio	Persen
X ₅	Angka Partisipasi Sekolah berusia 13 sampai 15 tahun	Rasio	Persen
X ₆	Angka Partisipasi Sekolah berusia 16 sampai 18 tahun	Rasio	Persen
X ₇	Sumber Penerangan Listrik	Rasio	Persen
X ₈	Tingkat Pengangguran Terbuka	Rasio	Persen

Tabel 3.1 Variabel Penelitian (Lanjutan)

Nama	Keterangan	Skala	Satuan
X_9	Tingkat Partisipasi Angkatan Kerja	Rasio	Persen
X_{10}	Tindak Pidana Terselesaikan	Rasio	Persen
X_{11}	Tingkat Keluhan Kesehatan	Rasio	Persen
X_{12}	Persentase Kemiskinan	Rasio	Persen
X_{13}	Rata-Rata Pengeluaran Perkapita Sebulan	Rasio	Ratus ribu Rupiah

Selanjutnya untuk struktur data akan ditampilkan pada Tabel 3.2 berikut

Tabel 3.2 Struktur Data dalam Penelitian

Wilayah	X_1	X_2	...	X_{13}
1	$x_{1.1}$	$x_{2.1}$...	$x_{13.1}$
2	$x_{1.2}$	$x_{2.2}$...	$x_{13.2}$
3	$x_{1.3}$	$x_{2.3}$...	$x_{13.3}$
\vdots	\vdots	\vdots	\ddots	\vdots
38	$x_{1.38}$	$x_{2.38}$...	$x_{13.38}$

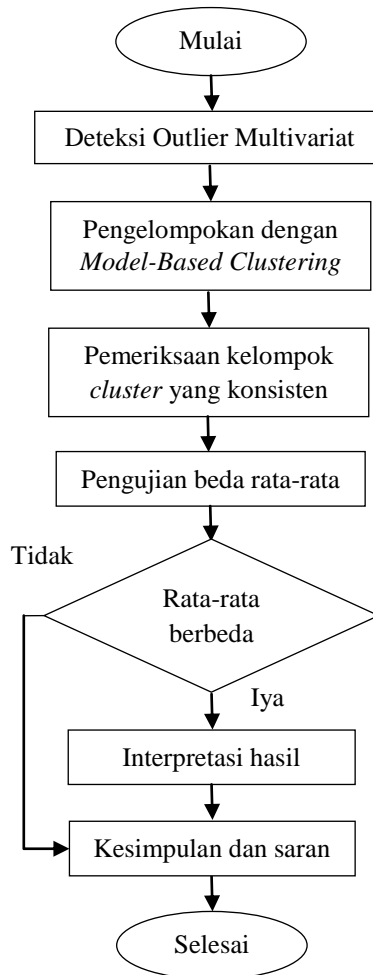
3.3 Langkah Penelitian

Adapun langkah analisis yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Melakukan deteksi outlier multivariat untuk mengetahui adanya data outlier secara multivariat
2. Melakukan pengelompokan objek penelitian dengan metode *model-based clustering* berdasarkan kriteria *Integrated Completed Likelihood*
3. Melakukan pemeriksaan untuk kelompok optimal yang paling konsisten
4. Melakukan pengujian perbedaan rata-rata menggunakan Manova
5. Melakukan interpretasi hasil pengelompokan jika setiap kelompok memiliki perbedaan nyata
6. Menarik kesimpulan dan saran

3.4 Diagram Alir

Berikut ini adalah diagram alir yang menyajikan proses penelitian secara keseluruhan



Gambar 3.1 Diagram Alir Penelitian

(Halaman ini sengaja dikosongkan)

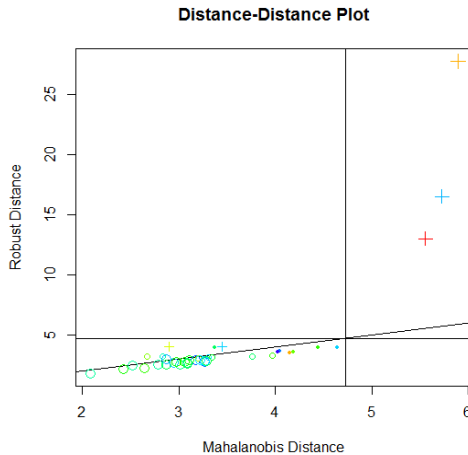
BAB IV

ANALISIS DAN PEMBAHASAN

Berdasarkan uraian pada latar belakang, tujuan utama penelitian ini adalah untuk mengelompokkan kabupaten/kota di Jawa Timur yang memiliki kemiripan karakteristik berdasarkan indikator-indikator pembangunan ekonomi dalam variabel penelitian. Proses pengelompokan dilakukan dengan metode *model-based clustering* distribusi t multivariat. Kriteria yang digunakan untuk memilih model terbaik adalah *Integrated Completed Likelihood* (ICL) yang dipilih dari nilai terbesar. Berikut adalah hasil dan pembahasan berdasarkan urutan yang sesuai dengan diagram alir penelitian

4.1 Deteksi *Outlier* Multivariat

Deteksi *outlier* multivariat pada data penelitian bertujuan untuk mendeteksi adanya *outlier* pada data indikator pembangunan ekonomi di Jawa Timur dari tahun ke tahun. Data *outlier* menyebabkan kurva menjadi lebih landai sehingga asumsi normal multivariat tidak terpenuhi dan penggunaan *model-based clustering* dengan distribusi t multivariat perlu dilakukan untuk mendapatkan hasil pengelompokan yang lebih baik. Jika dipaksakan menggunakan asumsi distribusi normal multivariat maka hasil penelitian akan bias dan tidak sesuai dengan kondisi nyata. Langkah yang digunakan dalam melakukan deteksi *outlier* multivariat pada data penelitian adalah dengan menghitung jarak mahalanobis dan jarak *robust*. Selanjutnya hasil tersebut digunakan akan dibandingkan dengan dengan kuantil dari distribusi $\chi^2_{p;0,975}$ sebagai nilai *cut off*. Nilai *cut off* yang terbentuk pada data setiap tahun akan berbeda-beda. Titik yang berada diluar batas kuantil akan ditandai dengan simbol berbeda sehingga terlihat bahwa kabupaten/kota tersebut adalah data *outlier*. Berikut adalah deteksi *outlier* multivariat secara visual menggunakan *Distance-distance plot*



Gambar 4.1 *Dd Plot* Data Indikator Pembangunan Ekonomi Tahun 2011

Gambar 4.1 merupakan plot jarak mahalanobis dan jarak *robust* pada data indikator pembangunan ekonomi pada tahun 2011. Berdasarkan Gambar 4.1 dapat diketahui bahwa terdapat tiga titik yang berada jauh dari garis lurus perpotongan *cut off* jarak mahalanobis dan jarak *robust*. Sedangkan titik-titik lain yang mewakili 35 kabupaten/kota cenderung berada disekitar garis diagonal. Hal tersebut menunjukkan terdapat tiga kabupaten/kota yang merupakan *outlier* data secara multivariat. Kabupaten/kota tersebut antara lain Kabupaten Sumenep, Kota Kediri dan Kota Surabaya .

Selain deteksi *outlier* yang dilakukan secara visual melalui *plot*, *package* yang digunakan juga memberikan output nilai jarak mahalanobis dan jarak *robust* sebagai bukti deteksi *outlier* secara perhitungan. Contoh hasil output deteksi outlier beserta nilai jarak mahalanobis dan jarak *robust* berada pada Lampiran 6. Pada *output* yang berisi “TRUE” dan “FALSE” keterangan “TRUE” menunjukkan kabupaten/kota yang terdeteksi sebagai *outlier* sedangkan keterangan “FALSE” menunjukkan bahwa kabupaten/kota tersebut tidak terdeteksi sebagai *outlier*. Berikut adalah hasil deteksi *outlier* data setiap tahun

Tabel 4.1 Daftar Kabupaten/Kota Yang Terdeteksi Sebagai Data Outlier

tahun	outlier	Kabupaten/Kota
2011	3	Sumenep, Kota Kediri, Kota Surabaya
2012	3	Sumenep, Kota Kediri, Kota Surabaya
2013	3	Sumenep, Kota Kediri, Kota Surabaya
2014	3	Sumenep, Kota Kediri, Kota Surabaya
2015	3	Sumenep, Kota Kediri, Kota Surabaya

Pada Tabel 4.1 ditunjukkan kabupaten/kota yang terdeteksi sebagai outlier. Hasil yang diperoleh berdasarkan pengamatan yang dilakukan tahun 2011 sampai dengan tahun 2015 diketahui bahwa jumlah outlier yang didapatkan tetap sebanyak tiga. kabupaten/kota yang termasuk anggota *outlier* juga tetap ditempati Kabupaten Sumenep, Kota Kediri dan Kota Surabaya. Hal tersebut menunjukkan bahwa tiga kabupaten/kota tersebut memiliki perbedaan yang cukup jauh dari kabupaten/kota yang lain jika ditinjau dari indikator pembangunan ekonomi setiap tahunnya. Kota Kediri dan Surabaya cenderung memiliki tingkat perekonomian yang sangat baik jika ditinjau dari Pendapatan Regional Domestik Bruto yang sangat tinggi. Hal ini dikarenakan Kota Surabaya merupakan pusat kegiatan perindustrian dan perekonomian di Jawa Timur. Kemudian Kota Kediri juga memiliki tingkat perekonomian yang tinggi karena keberadaan industri besar seperti industri rokok yang menyumbang nilai perekonomian sangat besar. Sedangkan Kabupaten Sumenep cenderung tertinggal jika dilihat dari beberapa variabel indikator perekonomian yang memiliki nilai rendah dari kabupaten/kota lain.

4.2 Pengelompokan Kabupaten/Kota Kriteria ICL

Pada penelitian ini data yang digunakan diasumsikan berdistribusi *t* multivariat sehingga penggunaan metode *model-based clustering* dapat dilakukan. Proses pengelompokan kabupaten/kota di Jawa Timur dengan metode *Model-Based Clustering* kriteria *Intregated Completed Likelihood* dilakukan

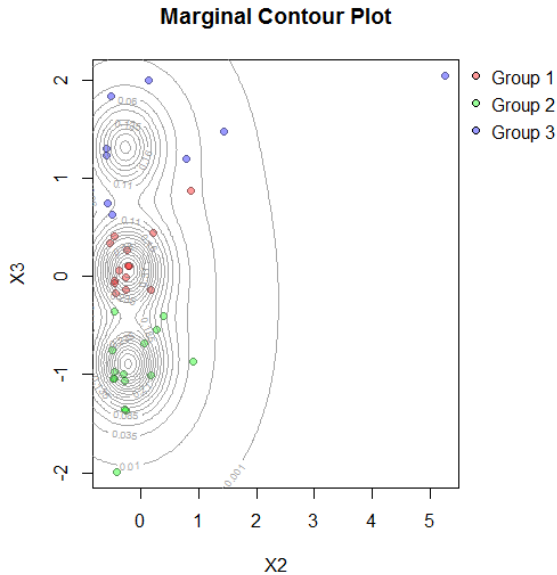
dengan *package teigen* pada *software R* Jumlah model yang mampu diidentifikasi adalah 28 model yang dengan jumlah kelompok maksimal 9 kelompok (Lampiran 11). Pemilihan model terbaik dan jumlah kelompok optimal berdasarkan kriteria ICL adalah dengan memilih model dengan nilai ICL terbesar diantara semua iterasi yang dilakukan. Berikut adalah hasil pengelompokkan kabupaten/kota di Jawa Timur mulai tahun 2011 sampai dengan tahun 2015

4.2.1 Pengelompokan Kabupaten/Kota tahun 2011

Pengolahan data Indikator pembangunan Ekonomi di Jawa Timur pada tahun 2011 untuk mengelompokkan kabupaten/kota di Jawa Timur menggunakan *package teigen* pada *software R* menghasilkan nilai ICL terbaik sebesar -1267.581. Meskipun iterasi dilakukan sampai pada jumlah 4 kelompok namun nilai ICL terbesar didapatkan pada jumlah kelompok sebanyak 3 *cluster*. Hal tersebut menunjukkan bahwa jumlah kelompok terbaik tidak selalu berada pada akhir iterasi. Adapun kerangka model yang didapatkan pada jumlah *cluster* optimal adalah model dengan kode CICC. Secara teori yang didasarkan pada tinjauan pustaka maka model CICC menunjukkan bentuk *cluster* sebagai berikut

- λ_g dengan label C (*Constrained*) berarti tiga kelompok yang terbentuk memiliki volume elips yang sama
- D_g dengan label I berarti matriks ortogonal *eigenvector* membentuk matriks Identitas
- A_g dengan label C (*Constrained*) berarti kontur yang terbentuk dari tiga kelompok sama
- D_g^T dengan label C (*Constrained*) berarti tiga kelompok yang terbentuk memiliki derajat bebas yang sama

Selanjutnya untuk mengetahui visualisasi hasil pengelompokan kabupaten/kota di Jawa Timur berdasarkan data Indikator Pembangunan Ekonomi pada tahun 2011, maka dilakukan *plotting* pada *marginal contour plot* seperti pada gambar berikut



Gambar 4.2 *Marginal Contour Plot* Data Tahun 2011

Gambar 4.2 menunjukkan visualisasi anggota kelompok yang terbentuk. Pemilihan variabel untuk visualisasi kontur dilakukan dengan mengkombinasikan 2 dari 13 variabel yang digunakan. Variabel X_2 merupakan PDRB atas harga berlaku dan X_3 adalah Indeks Pembangunan Manusia. Kedua variabel tersebut merupakan variabel yang paling berpengaruh pada pengelompokan tahun 2011 dimana kombinasi yang terbentuk selalu memberikan jumlah kontur yang sesuai. Visualisasi kontur yang terbentuk oleh kombinasi dua variabel ini adalah yang paling baik karena cenderung homogen pada setiap kelompoknya dan tidak terpotong dengan kelompok yang lain. Karakteristik kontur plot pada setiap kelompok memiliki bentuk dan volume yang sama. Titik-titik berwarna merah mewakili anggota kelompok pertama, titik berwarna hijau adalah kelompok kedua sedangkan titik berwarna biru merupakan anggota kelompok

ketiga. Untuk hasil pengelompokan dengan kriteria ICL secara lebih detail ditampilkan pada tabel berikut

Tabel 4.2 Hasil Pengelompokan Kabupaten/Kota Di Jawa Timur Tahun 2011

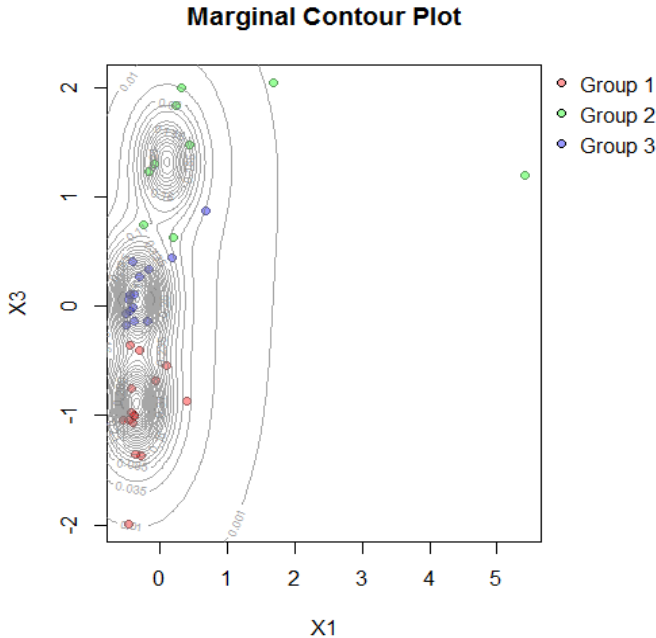
<i>Cluster 1</i>		<i>Cluster 2</i>		<i>Cluster 3</i>	
1	Sidoarjo	1	Pacitan	1	Ponorogo
2	Kota-kediri	2	Trenggalek	2	Tulungagung
3	Kota-Blitar	3	Malang	3	Blitar
4	Kota-Malang	4	Lumajang	4	Kediri
5	Kota-Pasuruan	5	Jember	5	Banyuwangi
6	Kota-Mojokerto	6	Bondowoso	6	Mojokerto
7	Kota-Madiun	7	Situbondo	7	Jombang
8	Kota-Surabaya	8	Probolinggo	8	Nganjuk
9	Kota-Batu	9	Pasuruan	9	Madiun
		10	Bojonegoro	10	Magetan
		11	Tuban	11	Ngawi
		12	Bangkalan	12	Lamongan
		13	Sampang	13	Gresik
		14	Pamekasan	14	Kota-Probolinggo
		15	Sumenep		

4.2.2 Pengelompokan Kabupaten/Kota tahun 2012

Hasil pengolahan data Indikator pembangunan Ekonomi di Jawa Timur pada tahun 2012 menunjukkan nilai ICL terbesar adalah -1222.373 Nilai tersebut didapatkan pada jumlah kelompok sebanyak 3 *cluster* dengan model CICU. Model tersebut menunjukkan bentuk *cluster* sebagai berikut

- λ_g dengan label C (*Constrained*) berarti tiga kelompok yang terbentuk memiliki volume elips yang sama
- D_g dengan label I berarti matriks ortogonal *eigenvector* membentuk matriks Identitas
- A_g dengan label C (*Constrained*) berarti kontur yang terbentuk dari tiga kelompok sama

- D_g^T dengan label U (*Unconstrained*) berarti tiga kelompok yang terbentuk memiliki derajat bebas yang berbeda
- Hasil *marginal contour plot* data Indikator Pembangunan Ekonomi pada tahun 2012 adalah sebagai berikut



Gambar 4.3 *Marginal Contour Plot* Data Tahun 2012

Berdasarkan Gambar 4.3 yang merupakan visualisasi anggota kelompok yang terbentuk ditunjukkan bahwa kontur plot pada setiap kelompok memiliki bentuk dan volume yang sama. Perbedaan yang terjadi dengan model pada tahun sebelumnya terdapat pada matrik *transpose* eigenvector (D_g^T) dengan label U (*Unconstrained*). Sehingga model tersebut akan memiliki elips dengan volume dan bentuk yang sama namun berbeda dalam derajat bebasnya. Pola kontur yang ditunjukkan juga tidak terlalu berbeda dengan tahun sebelumnya karena model yang terbentuk hanya berbeda pada matriks D_g^T . Selanjutnya perbedaan anggota

ditunjukkan dengan perbedaan warna yaitu merah untuk *cluster* 1, hijau untuk *cluster* 2 dan biru untuk *cluster* 3. Berikut adalah tabel rincian hasil pengelompokan kabupaten/kota di Jawa Timur pada tahun 2012

Tabel 4.3 Hasil pengelompokan Kabupaten/Kota di Jawa Timur tahun 2012

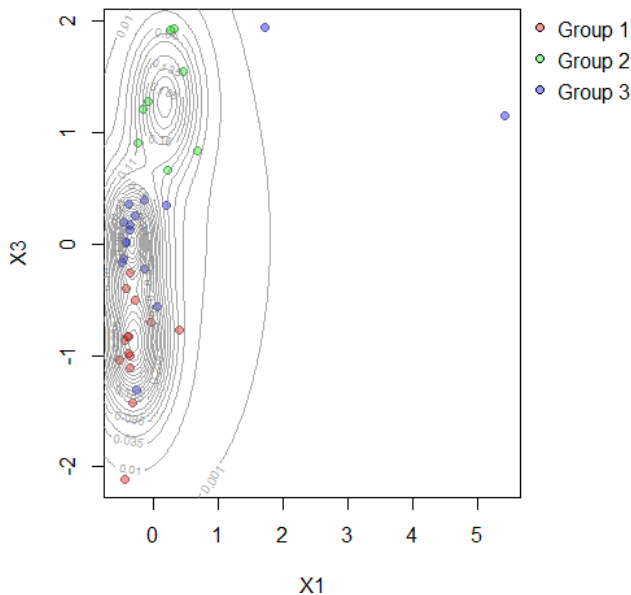
<i>Cluster 1</i>		<i>Cluster 2</i>		<i>Cluster 3</i>	
1	Sidoarjo	1	Ponorogo	1	Pacitan
2	Gresik	2	Trenggalek	2	Malang
3	Kota-Kediri	3	Tulungagung	3	Lumajang
4	Kota-Blitar	4	Blitar	4	Jember
5	Kota-Malang	5	Kediri	5	Bondowoso
6	Kota-Pasuruan	6	Banyuwangi	6	Situbondo
7	Kota-Mojokerto	7	Mojokerto	7	Probolinggo
8	Kota-Madiun	8	Jombang	8	Pasuruan
9	Kota-Surabaya	9	Nganjuk	9	Bojonegoro
10	Kota-Batu	10	Madiun	10	Bangkalan
		11	Magetan	11	Sampang
		12	Ngawi	12	Pamekasan
		13	Tuban	13	Sumenep
		14	Lamongan		
		15	Kota- Probolinggo		

4.2.3 Pengelompokan Kabupaten/Kota tahun 2013

Proses pengolahan data Indikator pembangunan Ekonomi di Jawa Timur pada tahun 2013 untuk mengelompokkan Kabupaten/Kota di Jawa Timur juga menghasilkan jumlah kelompok optimal sebanyak 3 *cluster*. Hasil pengelompokan tersebut didapatkan berdasarkan nilai ICL sebesar -1290.966 dengan kerangka model CICU. Kerangka model yang didapatkan sama dengan hasil pengelompokan Kabupaten/Kota di Jawa Timur pada tahun 2013 sehingga sifat dan bentuk kontur akan

sama. Sesuai penjelasan sebelumnya model CICU akan memiliki bentuk dan volume kontur yang sama antar kelompoknya. Selain itu derajat bebas antar kelompok juga memiliki kesamaan. Adapun *marginal contour plot* yang terbentuk pada analisis data Indikator pembangunan Ekonomi di Jawa Timur pada tahun 2013 akan disajikan pada gambar berikut

Marginal Contour Plot



Gambar 4.4 *Marginal Contour Plot* Data Tahun 2013

Pada Gambar 4.4 visualiasi anggota kelompok yang terbentuk tidak jauh berbeda dengan kontur pada tahun-tahun sebelumnya. Gambar yang terbentuk memiliki jumlah elips sebanyak 3 sesuai dengan jumlah kelompok yang terbentuk. Namun pada elips *cluster* pertama dengan tanda titik merah dan *cluster* ketiga bertanda titik biru terdapat irisan. Hal tersebut menunjukkan bahwa secara visual *cluster* tersebut tidak terlalu berbeda karakteristiknya. Kemudian rincian hasil pengelompokan kabupaten/kota di Jawa Timur tahun 2013 adalah sebagai berikut

Tabel 4.4 Hasil Pengelompokan Kabupaten/Kota Di Jawa Timur Tahun 2013

<i>Cluster 1</i>		<i>Cluster 2</i>		<i>Cluster 3</i>	
1	Ponorogo	1	Pacitan	1	Sidoarjo
2	Tulungagung	2	Trenggalek	2	Kota-Kediri
3	Kediri	3	Blitar	3	Kota-Blitar
4	Banyuwangi	4	Malang	4	Kota-Malang
5	Mojokerto	5	Lumajang	5	Kota-Mojokerto
6	Jombang	6	Jember	6	Kota-Madiun
7	Nganjuk	7	Bondowoso	7	Kota-Surabaya
8	Madiun	8	Situbondo		
9	Magetan	9	Probolinggo		
10	Ngawi	10	Pasuruan		
11	Bojonegoro	11	Tuban		
12	Lamongan	12	Bangkalan		
13	Gresik	13	Sampang		
14	Sumenep	14	Pamekasan		
15	Kota-Probolinggo				
16	Kota-Pasuruan				
17	Kota-Batu				

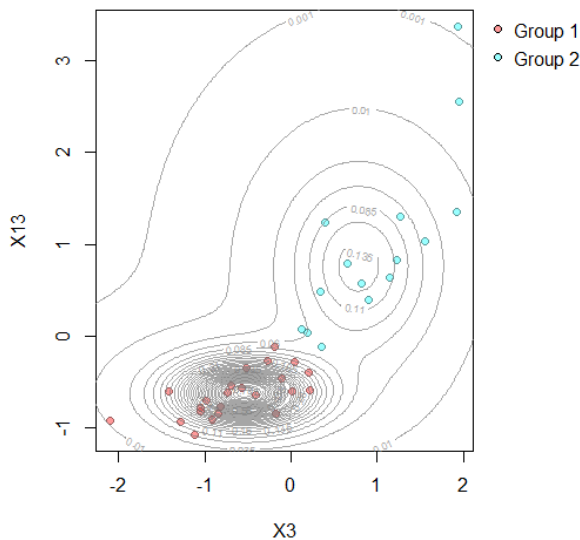
4.2.4 Pengelompokan Kabupaten/Kota tahun 2014

Pengelompokan kabupaten/kota di Jawa Timur berdasarkan data Indikator pembangunan Ekonomi di Jawa Timur pada tahun 2014 menunjukkan hasil yang berbeda dengan tahun-tahun sebelumnya. Jumlah kelompok optimal yang didapatkan pada tahun 2014 hanya 2 *cluster* saja meskipun iterasi maksimal berada pada 3 *cluster*. Hal tersebut terjadi karena Nilai ICL terbesar yang didapatkan berada pada $G=2$ yaitu dengan nilai -1229.807. Kemudian model kerangka untuk matriks adalah CIUC. Berdasarkan kerangka model pada tinjauan pustaka maka dugaan bentuk kontur yang ditunjukkan model CIUC adalah sebagai berikut

- λ_g dengan label C (*Constrained*) berarti tiga kelompok yang terbentuk memiliki volume elips yang sama
- D_g dengan label I berarti matriks ortogonal *eigenvector* membentuk matriks Identitas
- A_g dengan label U (*Unconstrained*) berarti bentuk kontur dari dua kelompok berbeda
- D_g^T dengan label C (*Constrained*) berarti dua kelompok yang terbentuk memiliki derajat bebas yang sama

Akibatnya pengelompokan kabupaten/kota di Jawa Timur berdasarkan data Indikator Pembangunan Ekonomi pada tahun 2014 memiliki *marginal contour plot* seperti pada gambar berikut

Marginal Contour Plot



Gambar 4.5 *Marginal Contour Plot* data tahun 2014

Gambar 4.5 menunjukkan visualisasi anggota kelompok yang terbentuk. Kontur plot pada setiap kelompok memiliki bentuk dan volume yang sama. Titik-titik berwarna merah mewakili anggota kelompok pertama, titik berwarna hijau adalah kelompok kedua sedangkan titik berwarna biru merupakan anggota kelompok

ketiga. Untuk hasil pengelompokan dengan kriteria ICL secara lebih detail ditampilkan pada tabel berikut

Tabel 4.5 Hasil pengelompokan Kabupaten/Kota di Jawa Timur tahun 2014

Cluster 1		Cluster 2	
1	Sidoarjo	1	Pacitan
2	Mojokerto	2	Ponorogo
3	Gresik	3	Trenggalek
4	Kota-Kediri	4	Tulungagung
5	Kota-Blitar	5	Blitar
6	Kota-Malang	6	Kediri
7	Kota-Probolinggo	7	Malang
8	Kota-Pasuruan	8	Lumajang
9	Kota-Mojokerto	9	Jember
10	Kota-Madiun	10	Banyuwangi
11	Kota-Surabaya	11	Bondowoso
12	Kota-Batu	12	Situbondo
		13	Probolinggo
		14	Pasuruan
		15	Jombang
		16	Nganjuk
		17	Madiun
		18	Magetan
		19	Ngawi
		20	Bojonegoro
		21	Tuban
		22	Lamongan
		23	Bangkalan
		24	Sampang
		25	Pamekasan
		26	Sumenep

4.2.5 Pengelompokan Kabupaten/Kota tahun 2015

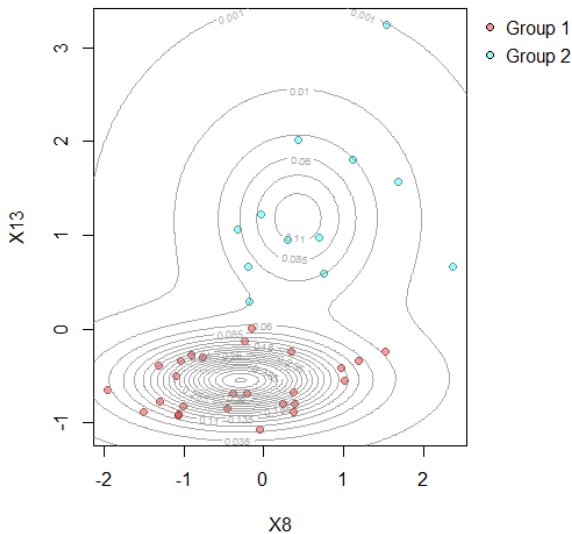
Pengolahan data Indikator pembangunan Ekonomi di Jawa Timur pada tahun 2015 untuk mengelompokkan kabupaten/kota di Jawa Timur menggunakan *package teigen* pada *software R* menghasilkan nilai ICL terbaik sebesar -1267.581. Meskipun iterasi dilakukan sampai pada jumlah 4 kelompok namun nilai ICL terbesar didapatkan pada jumlah kelompok sebanyak 3 *cluster* dengan model CICC. Berdasarkan kerangka model pada persamaan 2. maka model CICC menunjukkan bentuk *cluster* sebagai berikut

- λ_g dengan label C (*Constrained*) berarti tiga kelompok yang terbentuk memiliki volume elips yang sama
- D_g dengan label I berarti matriks ortogonal *eigenvector* membentuk matriks Identitas

- A_g dengan label C (*Constrained*) berarti kontur yang terbentuk dari tiga kelompok sama
- D_g^T dengan label C (*Constrained*) berarti tiga kelompok yang terbentuk memiliki derajat bebas yang sama

Akibatnya pengelompokan kabupaten/kota di Jawa Timur berdasarkan data Indikator Pembangunan Ekonomi pada tahun 2015 memiliki *marginal contour plot* seperti pada gambar berikut

Marginal Contour Plot



Gambar 4.6 *Marginal Contour Plot* Data Tahun 2015

Gambar 4.6 menunjukkan visualisasi anggota kelompok yang terbentuk. Secara teori kontur plot pada setiap kelompok memiliki bentuk dan volume yang sama. Hasil yang berbeda dapat dikarenakan sudut pandang pengambilan kontur dan kombinasi variabel yang berbeda pula. Titik-titik berwarna merah mewakili anggota kelompok pertama, titik berwarna hijau adalah kelompok kedua sedangkan titik berwarna biru merupakan anggota kelompok ketiga. Untuk hasil pengelompokan dengan kriteria ICL secara lebih detail ditampilkan pada tabel berikut

Tabel 4.6 Hasil pengelompokan Kabupaten/Kota di Jawa Timur tahun 2015

<i>Cluster 1</i>				<i>Cluster 2</i>	
1	Pacitan	14	Pasuruan	1	Sidoarjo
2	Ponorogo	15	Jombang	2	Mojokerto
3	Trenggalek	16	Nganjuk	3	Gresik
4	Tulungagung	17	Madiun	4	Kota-Kediri
5	Blitar	18	Magetan	5	Kota-Blitar
6	Kediri	19	Ngawi	6	Kota-Malang
7	Malang	20	Bojonegoro	7	Kota-Probolinggo
8	Lumajang	21	Tuban	8	Kota-Pasuruan
9	Jember	22	Lamongan	9	Kota-Mojokerto
10	Banyuwangi	23	Bangkalan	10	Kota-Madiun
11	Bondowoso	24	Sampang	11	Kota-Surabaya
12	Situbondo	25	Pamekasan	12	Kota-Batu
13	Probolinggo	26	Sumenep		

4.3 Uji Kesamaan Kelompok

Pengujian kesamaan kelompok adalah pengujian yang bertujuan untuk mengetahui adanya perbedaan secara signifikan antar kelompok yang terbentuk. Proses pengujian dilakukan merupakan uji beda rata-rata menggunakan uji Manova. Manova merupakan salah satu analisis multivariat yang digunakan untuk mengetahui apakah vektor rata-rata populasi sama. Adapun hipotesis yang digunakan untuk pengujian perbedaan rata-rata antar dua kelompok adalah sebagai berikut.

$H_0 : \tau_1 = \tau_2 = 0$ (tidak terdapat perbedaan antar kelompok)

$H_1 : \tau_1 \neq \tau_2$ (terdapat perbedaan antara kelompok 1 dan 2)

Sedangkan untuk pengujian dengan jumlah kelompok sebanyak tiga adalah sebagai berikut

$H_0 : \tau_1 = \tau_2 = \tau_3 = 0$ (tidak ada perbedaan antar kelompok)

$H_1 : \text{minimal satu } \tau_i \neq \tau_j$ (setidaknya terdapat satu kelompok ke-I dan kelompok ke-j yang berbeda)

Keputusan pengujian adalah tolak H_0 jika $p - value < \alpha$ pada taraf nyata 0.05. Berikut adalah hasil pengujian kesamaan kelompok yang terbentuk

Tabel 4.7 Hasil Uji Kesamaan Kelompok Dengan Nilai *Wald Type*

Tahun	Jumlah <i>cluster</i>	Wald Type	<i>p - value</i>
2011	3	3276.7380	0.000
2012	3	834.9354	0.000
2013	3	1551.6440	0.000
2014	2	224.1904	0.000
2015	2	194.9459	0.000

Hasil pengujian pada Tabel 4.7 menunjukkan bahwa keputusan pengujian kesamaan kelompok tahun 2011 sampai 2015 adalah tolak H_0 pada taraf nyata atau alfa sebesar 0.05. Hal tersebut ditunjukkan oleh *p-value* pengujian yang sangat kecil dan kurang dari nilai alfa. Karena pada tahun 2014 dan 2015 hasil pengelompokan adalah dua *cluster* maka kesimpulan pengujian adalah kelompok 1 berbeda secara nyata dengan kelompok kedua. Sedangkan pada tahun 2011, 2012 dan 2013 kesimpulan pengujian adalah minimal terdapat satu *cluster* ke-i yang memiliki perbedaan mean secara signifikan dengan *cluster* ke-j. Selanjutnya untuk mengetahui *cluster* mana yang memiliki perbedaan signifikan secara parsial maka dilakukan pengujian sebagai berikut

Tabel 4.8 Hasil Uji Kesamaan Kelompok Secara Parsial

Tahun	<i>Cluster</i>	Wald Type	<i>p - value</i>
2011	1 dan 2	554.3489	0.000
	2 dan 3	195.8184	0.000
	1 dan 3	592.2664	0.000
2012	1 dan 2	167.2995	0.000
	2 dan 3	111.3004	0.000
	1 dan 3	406.0304	0.000

Tabel 4.8 Hasil Uji Kesamaan Kelompok Secara Parsial (Lanjutan)

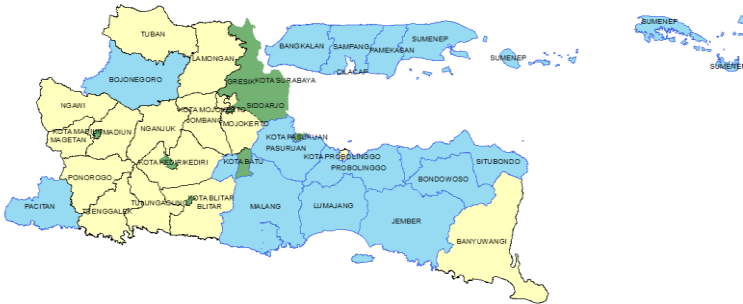
Tahun	Cluster	Wald Type	<i>p – value</i>
2013	1 dan 2	144.8273	0.000
	2 dan 3	786.3850	0.000
	1 dan 3	377.7437	0.000

Berdasarkan Tabel 4.8 hasil pengujian secara parsial menunjukkan bahwa *p-value* setiap *cluster* pada setiap tahun juga sangat kecil. Keputusan pengujian pada taraf nyata 0.05 adalah tolak H_0 sehingga kesimpulan yang didapatkan adalah antar kelompok pada setiap tahun memiliki perbedaan rata-rata yang signifikan. Dengan bukti bahwa setiap kelompok memiliki perbedaan vektor *mean* maka pengelompokan dapat dikatakan sudah berhasil.

4.4 Analisis Karakteristik Kelompok

Setelah diketahui bahwa kelompok yang terbentuk memiliki perbedaan *mean* maka perlu dilakukan analisis untuk mengetahui karakteristik antar kelompok yang terbentuk. Ketika uji menunjukkan hasil signifikan maka analisis karakteritik juga akan berguna sebagai pembeda tiap *cluster*. Untuk mempermudah pola pengelompokan yang terbentuk maka dilakukan visualisasi pada hasil pengelompokan berupa peta sebaran kelompok. Berikut ini adalah peta sebaran pengelompokkan kabupaten/kota setiap tahunnya

**Gambar 4.7** Peta Sebaran Pengelompokkan Tahun 2011



Gambar 4.8 Peta Sebaran Pengelompokan Tahun 2012



Gambar 4.9 Peta Sebaran Pengelompokan Tahun 2013

Berdasarkan peta sebaran yang terbentuk terlihat bahwa pengelompokan kabupaten/kota tahun 2011 sampai tahun 2013 selalu menghasilkan 3 kelompok. Gambar menunjukkan kelompok dengan warna hijau beranggotakan Kabupaten Sidoarjo dan semua kota kecuali Kota Probolinggo. Jika dilihat dari nilai PDRB perkapita maka *cluster* tersebut merupakan *cluster* dengan PDRB perkapita yang tinggi. Hal ini karena anggota kelompok pertama merupakan kota-kota yang maju seperti Kota Surabaya yang merupakan pusat perekonomian dan industri di Jawa Timur. Sedangkan Kota Kediri memiliki tingkat perekonomian tinggi karena memiliki industri besar seperti pabrik rokok. Anggota kelompok yang konsisten pada *cluster* tersebut sampai tahun 2013 adalah Kabupaten Sidoarjo, Kota Kediri, Kota Blitar, Kota Malang, Kota Mojokerto, Kota Madiun dan Kota Surabaya. Pada

tahun 2012 anggota *cluster* pertama bertambah dengan Kabupaten Gresik yang sebelumnya berada pada kelompok PDRB perkapita sedang. Pada tahun selanjutnya Kota Gresik, Malang dan Batu berpindah ke *cluster* dengan PDRB perkapita sedang. Kabupaten Gresik tidak konsisten pada *clusternya* karena inflasi pada tahun tersebut cenderung naik turun. Nilai inflasi Kabupaten Gresik terendah berada pada tahun 2012.

Cluster kedua merupakan *cluster* dengan tingkat PDRB perkapita rendah. Pada Gambar 4.7 *cluster* ini diberi warna biru dengan mayoritas anggota adalah Kabupaten yang dekat dengan pantai. Anggota yang konsisten pada *cluster* ini adalah Kabupaten Pacitan, Malang, Lumajang, Jember, Bondowoso, Situbondo, Probolinggo, Pasuruan, Bangkalan, Sampang dan Pamekasan. Anggota konsisten pada *cluster* tersebut memiliki karakteristik ekonomi yang rendah contohnya Kabupaten Sampang dengan persentase kemiskinan yang sangat tinggi. Kabupaten Tuban dan Pacitan merupakan contoh kabupaten yang tidak konsisten. Dua kabupaten tersebut berpindah *cluster* dari kelompok rendah ke kelompok sedang pada tahun 2012 dan kembali ke kelompok rendah pada tahun 2013. Hal ini terjadi karena sektor industri dan pertanian memuncak di tahun 2012 untuk dua kabupaten tersebut yang menyebabkan indikator perekonomian pada tahun tersebut membaik.

Selanjutnya *cluster* ketiga dengan warna kuning merupakan kelompok dengan tingkat PDRB sedang. Anggota konsisten pada *cluster* ini adalah Kabupaten Ponorogo, Tulungagung, Banyuwangi, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Lamongan dan Kota Probolinggo. Anggota pada klaser ini merupakan kabupaten/kota yang sedang berkembang misalnya Kabupaten Banyuwangi yang terus mengembangkan potensi ekonomi melalui sektor budaya. Kemudian Kabupaten Mojokerto yang terus membenahi perekonomian melalui sektor industri. Berdasarkan perbandingan hasil pengelompokkan tersebut maka pada periode tahun 2011 sampai tahun 2013 terdapat sebanyak 29

kabupaten/kota yang konsisten atau tidak mengalami perubahan *cluster*



Gambar 4.10 Peta Sebaran Pengelompokkan Tahun 2014



Gambar 4.11 Peta Sebaran Pengelompokkan Tahun 2015

Berbeda dengan tahun-tahun sebelumnya, pengelompokan pada tahun 2014 dan 2015 hanya menghasilkan kelompok optimal 2 *cluster*. Hasil pengelompokan pada tahun 2014 dan 2015 tidak memiliki perbedaan anggota kelompok sehingga dapat dikatakan pada dua tahun tersebut kelompok yang terbentuk sudah konsisten. Pola pengelompokan yang terbentuk hampir sama dengan tahun-tahun sebelumnya yaitu *cluster* pertama beranggotakan semua Kota ditambah Kabupaten Sidoarjo, Mojokerto dan Gresik. Sedangkan Kabupaten lain yang

sebelumnya cenderung berada pada tingkat PDRB perkapita sedang dan rendah bergabung pada *cluster* kedua. Untuk perbandingan nilai *mean* pada hasil pengelompokan ditampilkan pada tabel berikut

Tabel 4.9 Nilai *Mean* pada Hasil Pengelompokkan Tahun 2011-2015

var	clas	Tahun				
		2011	2012	2013	2014	2015
x1	1	63.1846	69.3816	27.2662	68.2597	30.2965
	2	19.0626	19.9926	21.6436	28.0026	75.5695
	3	20.3955	20.6020	88.9995		
x2	1	53.1779	61.2250	25.2853	76.6030	25.1625
	2	23.2867	19.8949	26.9614	23.4473	84.2061
	3	20.9971	26.2368	83.5114		
x3	1	73.9544	74.4580	68.5447	74.7508	66.1858
	2	60.8833	67.1667	62.9407	65.4292	75.4558
	3	67.0543	61.4846	76.4814		
x13	1	7.1682	7.4548	5.7954	9.3049	6.6647
	2	3.6988	4.7623	4.6094	5.5257	11.7438
	3	4.5454	3.9530	8.7139		

Tabel 4.9 merupakan tabel perbedaan mean variabel yang dianggap memiliki pengaruh terbesar. Tabel secara lengkap dapat dilihat pada lampiran 18. Hasil pengelompokkan pada tahun 2015 menunjukkan bahwa kelompok 1 merupakan kelompok kabupaten/kota di Jawa Timur dengan karakteristik nilai PDRB, Angka Partisipasi sekolah dan konsumsi pengeluaran rumah tangga yang tinggi. Sedangkan pada kelompok 2 maupun 3 pada hasil pengelompokan 3 *cluster* adalah Kabupaten/Kota di Jawa Timur dengan karakteristik nilai PDRB, Angka Partisipasi sekolah dan konsumsi pengeluaran rumah tangga menengah ke bawah. Hal tersebut berlaku pada setiap tahun tidak terkecuali pada tahun 2014 dan 2015 yang hanya menghasilkan 2 *cluster*. Secara umum *cluster* kedua pada tahun tersebut merupakan gabungan *cluster* 2 dan 3 pada tahun sebelumnya. Hasil karakteristik yang cenderung sama juga didapatkan untuk pengelompokkan data pada tahun-tahun sebelumnya.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan dapat diambil kesimpulan sebagai berikut:

1. Hasil deteksi *outlier* menggunakan *distance-distance plot* menunjukkan bahwa pada kuantil 90 persen data yang digunakan dalam penelitian ini mengandung *outlier* sebanyak 3 pada setiap tahunnya.
2. Pengelompokan menggunakan *model-based clustering* dengan kriteria *Integrated Completed Likelihood* (ICL) menunjukkan bahwa jumlah kelompok optimal yang terbentuk pada tahun 2011, 2012 dan 2013 adalah 3 klaster. Sedangkan pada tahun 2014 dan 2015 terbentuk 2 klaster optimal.
3. Hasil pengelompokan pada tahun 2011 sampai dengan 2013 menunjukkan bahwa mayoritas kelompok konsisten berada pada klaster kedua dan ketiga dengan jumlah anggota konsisten masing-masing 11. Kemudian pada tahun 2014 dan 2015 kelompok konsisten mayoritas berada pada klaster kedua dengan jumlah 22 Kabupaten yang dapat dikatakan sebagai gabungan dua klaster sebelumnya. Sehingga secara keseluruhan karakteristik pengelompokan yang dilakukan cenderung memiliki kesamaan yaitu klaster pertama beranggotakan kabupaten/kota dengan indikator perekonomian yang maju. Sedangkan kabupaten/kota dengan tingkat perekonomian menengah ke bawah berada klaster selanjutnya.

5.2 Saran

Dari hasil penelitian menunjukkan bahwa mayoritas Kabupaten/Kota di Jawa Timur berada pada tingkat ekonomi menengah ke bawah. Saran bagi pemerintah adalah diharapkan

lebih memperhatikan kabupaten/kota pada kelompok kedua sehingga kebijakan-kebijakan yang diberikan pada kelompok tersebut mampu meningkatkan pertumbuhan ekonomi. Indikator indikator yang perlu diperhatikan antara lain PDRB regional, Tingkat pendidikan serta kualitas ekonomi setiap rumah tangga. Untuk penelitian selanjutnya disarankan untuk lebih memperhatikan distribusi data yang akan digunakan sebagai dasar pengelompokan pada *model-based clustering*. Selain itu metode diharapkan menggunakan pengujian asumsi distribusi *t* multivariat yang lebih *robust* sehingga hasil uji distribusi lebih tepat dan penelitian yang didapatkan juga lebih baik.

DAFTAR PUSTAKA

- Agustini, M. (2017). *Model-Based Clustering Dengan Distribusi tMultivariat Menggunakan Kriteria Integrated Completed Likelihood dan Minimum Message Length*. Tesis. Departemen Statistika Institut Teknologi Sepuluh Nopember.
- Andrews, J.L., McNicholas, P.D. & Subedi, S. (2011). Model-based classification via mixtures of multivariate t distributions. *Computational Statistics and Data Analysis* 55(1), 520-529.
- Ardani, A. (1992). *Analysis of Regional Growth and Disparity: the Impact Analysis of The Project on Indonesian Development*. Unpublished Dissertation. University of Pennsylvania Philadelphia.
- Arsyad, L. (1999). *Pengantar Perencanaan dan Pembangunan Ekonomi Daerah*. Yogyakarta: BPFE.
- Arsyad, L. (2010). *Ekonomi Pembangunan*. Yogyakarta: UPP STIM YKPN.
- Badan Pusat Statistik Provinsi Jawa Timur. (2017). *Berita Resmi Statistik No.16/02/Th.XX, 06 Februari 2017*. Surabaya: BPS Provinsi Jawa Timur.
- Banfield, J.D. & Raftery, A.E.(1993). Model-Based Gaussian and non-Gaussian Clustering. *Biometrics*, 49(3),803-821.
- Bathke, A., Friedrich, S., Konietzschke, F., Pauly, M., Strobl, N. & Hoeller, Y. (2016). Using EEG, SPECT, and Multivariate Resampling Methods to Differentiate Between Alzheimer's and other Cognitive Impairments. *arXiv preprint arXiv*, DOI: 1606.09004
- Baudry, J.P., Cardoso, M., Celeux, G., Amorim, M.J. & Ferreira, A.S. (2013). Enhancing the selection of a model-based clustering with external categorical variables. *Advances in Data Analysis and Classification*, 1-20. DOI: 1862-5347.
- Biernacki, C., Celeux, G. & Govaert, G. (2000). Assessing a Mixture Model for Clustering with the Integrated

- Completed Likelihood. *IEEE Trans. Pattern Anal. Mach. Intel.* 719-725.
- Bouveyron, C. & Brunet-Saumard, C. (2014). Model-based clustering of high dimensional data: A review. *Computational Statistics and Data Analysis, Journal of the American Statistical Association*, 71, 52-78.
- Casella, G. & Berger, R.L. (2002). *Statistical Inference, Second Editon*. Pacific Grove: Thomson Learning Inc.
- Dekking, F.M., Kraaikamp, C., Lopuha, H.P., & Mester, L.E. (2005). *A Modern introduction to Probability and Statistics*. London: Springer.
- Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977). Maximum Likelihood for Incomplete Data via the EM Algorithm (with discussion). *Journal of the Royal Statistical Society*, 39, 1-38.
- Dumairy. (1996). *Perekonomian Indonesia*. Jakarta: Penerbit Erlangga.
- Efriza, U. (2014). *Analisis Kesenjangan Pendapatan Antar Kabupaten/Kota di Provinsi Jawa Timur di Era Desentralisasi Fiskal*. Jurusan Ilmu Ekonomi. Malang: Universitas Brawijaya.
- Fraley, C. & Raftery, A.E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97, 611-631.
- Hair, J.F., Black, W.C., Babin, B.J. & Anderson, R.E. (2010). *Multivariate Data Analysis*. New Jersey: Upper saddle river.
- Haryono, C.A. (2016). *Penerapan Metode C-Means dan Fuzzy C-Means pada Pengelompokan Kabupaten/Kota Di Provinsi Jawa Timur Berdasarkan Indikator Pembangunan Ekonomi*. Tugas Akhir. Departemen Statistika Institut Teknologi Sepuluh Nopember.
- Hasan, I. (2001). *Pokok - Pokok Materi Statistik 2*. Jakarta: Bumi Aksara.

- Hammouda, K.& Karray, F. (2000). A comparative study of data clustering techiques. *Course Project SYDE 625* (hal. Tools of Intelligent Systems Design). Ontario: Univ. of Waterloo.
- Huang, Z & Michael, K.N. (1999). A fuzzy k-modes algorithm for clustering categorical data. *IEEE Trans. Fuzzy Syst*, 7(4), 446–452.
- Irawan & Suparmoko, M. (1987). *Ekonomi Pembangunan*, Yogyakarta: Liberty.
- Jain, A. (2010). Data clustering: 50 years beyond K-mean. *Pattern Recognition Letters*, 31(8), 651-666.
- Jhingan, M.L. (1993). *Ekonomi Pembangunan dan Perencanaan*. Jakarta: Raja Grafindo Persada.
- Johnson, R.A.& Wichern, D.W. (2007). *Applied Multivariate Analysis*. Amerika: Pearson Education.
- Kuncoro, M. (2006). *Ekonomika Pembangunan Teori, Masalah dan Kebijakan*. Yogyakarta: UPP STIM YKPN.
- Mardiana, S.A. (2012). *Kondisi Ketimpangan Ekonomi Antar Kabupaten/Kota dan Implikasinya Terhadap Kebijakan Pembangunan di Provinsi Jawa Timur*. Department of Economic Science. Bogor: Institut Pertanian Bogor.
- McLachlan, G.J. & Peel, D. (2000). *Finite Mixture Models*. New York: JohnWiley and Sons.
- Morrison, D.F.(2005). *Multivariate Statistical Methods 4th edition*, United States of America.
- Pardede, T. (2008). Perbandingan metode berbasis model (model-based) dengan metode-metode K-Means dalam analisis gugus. *Jurnal Sigma, Sains danTeknologi* 11(2), 157-166.
- Rousseeuw, P.J. & Van Driessen, K. (1999). A Fast Algorithm for the Minimum Covariance Determinant Estimator. *Technometrics* 41, 212-223.
- Sharma, S. (1996). *Applied Multivariate Techniques*. New York: John Wiley and Sons, Inc.
- Simamora,B. (2002). *Analysis Multivariat Pemasaran*. Jakarta: PT. Gramedia Pustaka Utama.
- Sukirno, S. (1996). *Makro Ekonomi*. Jakarta: Erlangga.

- Sumitro, D. (1987). *Dasar Teori Ekonomi Pertumbuhan dan Pembangunan*. Jakarta: LP3ES.
- Susilawati, U. (2011). *Penerapan Metode Penggerombolan berdasarkan Gaussian Mixture Models dengan Menggunakan Algoritma Expectation Maximization* Tugas Akhir. Departemen Statistika Institut Pertanian Bogor.
- Syafrizal, H.S. (2008). *Ekonomi Regional: Teori dan Aplikasi*. Padang: Badouse Media.
- Todaro, M.P. dan Smith, S.C. (2006). *Pembangunan Ekonomi. Jilid I Edisi Kesembilan*. Terjemahan Haris Munandar. Jakarta: Erlangga.
- Todaro, M.P. (2010). *Economics Development*. New York: Addison Wesley.
- Walpole, R.E. (2007). *Pengantar Statistika. Edisi ke- 3*. Terjemahan Bambang Sumantri. Jakarta: PT. Gramedia Pustaka Utama.

BIODATA PENULIS



Penulis bernama Shiddiqul Akhyar lahir di Lumajang pada tanggal 13 Januari 1996, anak pertama dari pasangan Mokhammad Salim dan Nuryani. Penulis bertempat tinggal di Desa Wonoayu, Kecamatan Ranuyoso, Kabupaten Lumajang, Jawa Timur. Pendidikan formal yang telah ditempuh dimulai dari SDN Wonoayu 02 (2001-2005), SDN Ranuyoso 01 (2005-2007), SMPN 1 Ranuyoso (2007-2010) dan SMAN 2 Lumajang (2010-2013). Kemudian penulis menempuh pendidikan Sarjana di Institut Teknologi Sepuluh Nopember Surabaya program studi S1 Statistika dengan NRP 1313 100 068. Pada masa perkuliahan, penulis pernah melakukan Kerja Praktek di Balai Penelitian Tanaman Jeruk dan Buah Subtropika Kota Batu selama satu bulan. Kepanitiaan yang pernah diikuti adalah *Creative Writing Talk* dari divisi PERS Himpunan Mahasiswa Statistika dan acara Pekan Raya Statistika 2015. Selain itu penulis menyalurkan hobi dalam musik dengan bergabung pada band perwakilan himpunan mahasiswa bernama *Sigma Project*. Dari pengalaman tersebut penulis banyak sekali belajar berorganisasi, manajemen waktu dan belajar menerapkan ilmu yang selama ini diperoleh di dalam perkuliahan untuk diterapkan pada permasalahan nyata. Segala kritik dan saran akan diterima oleh penulis untuk perbaikan kedepannya. Jika ada keperluan atau ingin berdiskusi dengan penulis mengenai Tugas Akhir ini dapat dihubungi melalui email shiddiqul.akhyar@gmail.com.

LAMPIRAN

Lampiran 1 Data Indikator Pembangunan Ekonomi Tahun 2011

No	Kabupaten/Kota	X_1	X_2	X_3	...	X_{13}
1	Pacitan	13.9586	7.5921	62.03	...	3.8159
2	Ponorogo	11.5912	9.9603	65.28	...	3.7037
3	Trenggalek	13.1765	8.9441	64.27	...	3.88067
4	Tulungagung	18.8839	18.8595	67.76	...	4.30762
5	Blitar	16.0151	18.0134	65.47	...	4.29015
6	Kediri	13.534	20.4717	66.84	...	4.01867
7	Malang	19.0032	46.9757	63.97	...	4.43119
8	Lumajang	15.8658	16.0785	60.72	...	3.71314
9	Jember	15.7922	37.1595	60.64	...	3.73049
10	Banyuwangi	23.5794	36.9509	65.48	...	4.60379
11	Bondowoso	12.8506	9.5528	60.46	...	3.7454
12	Situbondo	14.5904	9.5364	60.82	...	3.83114
13	Probolinggo	15.2308	16.8745	60.3	...	3.67572
14	Pasuruan	45.4484	69.5491	61.43	...	3.89161
15	Sidoarjo	47.0091	93.2317	74.48	...	6.96468
16	Mojokerto	37.2297	38.6928	68.71	...	4.86258
17	Jombang	16.0545	19.4722	66.84	...	5.14167
18	Nganjuk	12.4152	12.7145	66.58	...	4.20894
19	Madiun	13.6847	9.1182	65.98	...	4.40314
20	Magetan	14.8193	9.231	68.52	...	4.4281
21	Ngawi	11.6166	9.5359	65.84	...	3.85525
22	Bojonegoro	33.9612	41.3571	63.22	...	3.72946
23	Tuban	27.8933	31.4604	62.47	...	3.97595
24	Lamongan	15.442	18.265	66.21	...	4.56808
25	Gresik	56.2446	67.2976	71.11	...	5.45658

Lampiran 1 Data Indikator Pembangunan Ekonomi Tahun 2011
(Lanjutan)

No	Kabupaten/Kota	X_1	X_2	X_3	...	X_{13}
26	Bangkalan	19.2797	17.7144	58.63	...	3.52982
27	Sampang	12.4607	11.1187	55.17	...	3.26054
28	Pamekasan	9.8122	7.9294	60.42	...	3.13193
29	Sumenep	16.615	17.4571	58.7	...	3.13891
30	Kota-Kediri	235.783	64.0177	72.93	...	6.21491
31	Kota-Blitar	23.8443	3.183	73.08	...	6.27224
32	Kota-Malang	42.2068	34.968	77.36	...	7.88192
33	Kota- Probolinggo	24.4273	5.3765	68.14	...	5.78749
34	Kota-Pasuruan	21.1711	3.9889	70.41	...	7.66782
35	Kota-Mojokerto	27.2521	3.3116	73.47	...	7.32541
36	Kota-Madiun	39.5961	6.8134	76.48	...	6.98965
37	Kota-Surabaya	93.8611	261.7723	77.62	...	9.38706
38	Kota-Batu	37.9382	7.315	69.76	...	5.81036

Lampiran 2 Data Indikator Pembangunan Ekonomi Tahun 2012

No	Kabupaten/Kota	X_1	X_2	X_3	...	X_{13}
1	Pacitan	13.9586	7.5921	62.03	...	3.8159
2	Ponorogo	11.5912	9.9603	65.28	...	3.7037
3	Trenggalek	13.1765	8.9441	64.27	...	3.88067
4	Tulungagung	18.8839	18.8595	67.76	...	4.30762
5	Blitar	16.0151	18.0134	65.47	...	4.29015
6	Kediri	13.534	20.4717	66.84	...	4.01867
7	Malang	19.0032	46.9757	63.97	...	4.43119
8	Lumajang	15.8658	16.0785	60.72	...	3.71314
9	Jember	15.7922	37.1595	60.64	...	3.73049
10	Banyuwangi	23.5794	36.9509	65.48	...	4.60379
11	Bondowoso	12.8506	9.5528	60.46	...	3.7454
12	Situbondo	14.5904	9.5364	60.82	...	3.83114
13	Probolinggo	15.2308	16.8745	60.3	...	3.67572
14	Pasuruan	45.4484	69.5491	61.43	...	3.89161
15	Sidoarjo	47.0091	93.2317	74.48	...	6.96468
16	Mojokerto	37.2297	38.6928	68.71	...	4.86258
17	Jombang	16.0545	19.4722	66.84	...	5.14167
18	Nganjuk	12.4152	12.7145	66.58	...	4.20894
19	Madiun	13.6847	9.1182	65.98	...	4.40314
20	Magetan	14.8193	9.231	68.52	...	4.4281
21	Ngawi	11.6166	9.5359	65.84	...	3.85525
22	Bojonegoro	33.9612	41.3571	63.22	...	3.72946
23	Tuban	27.8933	31.4604	62.47	...	3.97595
24	Lamongan	15.442	18.265	66.21	...	4.56808
25	Gresik	56.2446	67.2976	71.11	...	5.45658
26	Bangkalan	19.2797	17.7144	58.63	...	3.52982

Lampiran 2 Data Indikator Pembangunan Ekonomi Tahun 2012
(Lanjutan)

No	Kabupaten/Kota	X_1	X_2	X_3	...	X_{13}
27	Sampang	12.4607	11.1187	55.17	...	3.26054
28	Pamekasan	9.8122	7.9294	60.42	...	3.13193
29	Sumenep	16.615	17.4571	58.7	...	3.13891
30	Kota-Kediri	235.783	64.0177	72.93	...	6.21491
31	Kota-Blitar	23.8443	3.183	73.08	...	6.27224
32	Kota-Malang	42.2068	34.968	77.36	...	7.88192
33	Kota- Probolinggo	24.4273	5.3765	68.14	...	5.78749
34	Kota-Pasuruan	21.1711	3.9889	70.41	...	7.66782
35	Kota-Mojokerto	27.2521	3.3116	73.47	...	7.32541
36	Kota-Madiun	39.5961	6.8134	76.48	...	6.98965
37	Kota-Surabaya	93.8611	261.7723	77.62	...	9.38706
38	Kota-Batu	37.9382	7.315	69.76	...	5.81036

Lampiran 3 Data Indikator Pembangunan Ekonomi Tahun 2013

No	Kabupaten/Kota	X_1	X_2	X_3	...	X_{13}
1	Pacitan	17.1857	9.4163	63.38	...	4.46696
2	Ponorogo	14.0685	12.1536	67.03	...	4.84752
3	Trenggalek	16.0983	11.0079	65.76	...	5.1406
4	Tulungagung	23.0384	23.2553	69.3	...	5.70023
5	Blitar	19.1389	21.7552	66.49	...	5.30668
6	Kediri	16.4019	25.1032	68.01	...	4.98194
7	Malang	23.3885	58.6747	65.2	...	5.4798
8	Lumajang	19.1801	19.637	61.87	...	3.89223
9	Jember	18.9198	45.0555	62.43	...	4.40143
10	Banyuwangi	29.9287	47.3647	66.74	...	6.00861
11	Bondowoso	15.6652	11.7926	63.21	...	4.26717
12	Situbondo	18.0306	11.9128	63.43	...	4.53726
13	Probolinggo	18.5459	20.8309	62.61	...	5.14801
14	Pasuruan	54.5654	84.9426	63.74	...	4.8125
15	Sidoarjo	57.4642	117.7433	76.39	...	8.70342
16	Mojokerto	45.1465	47.7563	69.84	...	6.61564
17	Jombang	19.36	23.8298	68.63	...	5.18742
18	Nganjuk	15.1167	15.6245	68.98	...	5.11063
19	Madiun	16.8088	11.2936	68.07	...	5.29828
20	Magetan	18.0803	11.3129	69.86	...	5.60215
21	Ngawi	14.5845	12.0262	67.25	...	4.81569
22	Bojonegoro	39.2026	48.1292	64.85	...	5.0464
23	Tuban	34.173	39.0084	64.14	...	5.14853
24	Lamongan	19.3971	23.0123	68.9	...	6.34616
25	Gresik	67.7638	83.153	72.47	...	7.16215
26	Bangkalan	20.841	19.5384	60.19	...	4.14166

Lampiran 3 Data Indikator Pembangunan Ekonomi Tahun 2013
(Lanjutan)

No	Kabupaten/Kota	X_1	X_2	X_3	...	X_{13}
27	Sampang	15.3013	13.9777	56.45	...	3.97907
28	Pamekasan	11.9769	9.9098	62.27	...	3.81004
29	Sumenep	23.8973	25.36	60.84	...	3.73185
30	Kota-Kediri	288.6962	79.8589	74.18	...	6.90815
31	Kota-Blitar	28.9569	3.9295	74.53	...	7.38373
32	Kota-Malang	50.9274	42.8199	78.44	...	9.63162
33	Kota- Probolinggo	29.319	6.564	70.05	...	7.52278
34	Kota-Pasuruan	25.1356	4.8332	72.89	...	7.24104
35	Kota-Mojokerto	32.6005	4.0361	74.91	...	7.60324
36	Kota-Madiun	48.1889	8.3904	78.41	...	10.3461
37	Kota-Surabaya	116.1624	327.802	78.51	...	10.4209
38	Kota-Batu	46.2749	9.0786	71.55	...	7.30268

Lampiran 4 Data Indikator Pembangunan Ekonomi Tahun 2014

No	Kabupaten/Kota	X ₁	X ₂	X ₃	...	X ₁₃
1	Pacitan	10.4922	19.0947	63.81	...	4.92782
2	Ponorogo	13.3936	15.4695	67.4	...	4.93231
3	Trenggalek	12.298	17.9067	66.16	...	5.36635
4	Tulungagung	25.7806	25.3752	69.49	...	5.87981
5	Blitar	24.1406	21.1613	66.88	...	6.14103
6	Kediri	27.7538	18.0345	68.44	...	5.44672
7	Malang	65.9495	26.0971	65.59	...	5.99162
8	Lumajang	21.9832	21.4182	62.33	...	4.44452
9	Jember	50.5519	21.1107	62.64	...	4.97592
10	Banyuwangi	53.4063	33.6294	67.31	...	6.47852
11	Bondowoso	13.0741	17.2711	63.43	...	4.79905
12	Situbondo	13.3388	20.0279	63.91	...	5.07539
13	Probolinggo	23.1551	20.4425	63.04	...	5.22264
14	Pasuruan	94.905	60.468	64.35	...	5.41086
15	Sidoarjo	131.6453	63.1718	76.78	...	8.93096
16	Mojokerto	53.2026	49.6995	70.22	...	7.76123
17	Jombang	26.3391	21.3358	69.07	...	6.86798
18	Nganjuk	17.2485	16.6215	69.59	...	5.46765
19	Madiun	12.5317	18.5934	68.6	...	6.13002
20	Magetan	12.5647	20.0518	70.29	...	6.48152
21	Ngawi	13.3115	16.08	67.78	...	5.74794
22	Bojonegoro	50.5426	41.012	65.27	...	5.52953
23	Tuban	43.8538	38.2302	64.58	...	5.58975
24	Lamongan	25.7241	21.67	69.42	...	6.8011
25	Gresik	93.7851	75.5349	72.84	...	7.9363
26	Bangkalan	21.7092	22.9527	60.71	...	5.42979

Lampiran 4 Data Indikator Pembangunan Ekonomi Tahun 2014
(Lanjutan)

No	Kabupaten/Kota	X_1	X_2	X_3	...	X_{13}
27	Sampang	14.6283	15.7988	56.98	...	4.74872
28	Pamekasan	11.0785	13.2482	62.66	...	5.04876
29	Sumenep	28.3127	26.5298	61.43	...	4.73187
30	Kota-Kediri	87.7042	315.4009	74.62	...	8.07148
31	Kota-Blitar	4.3541	31.804	75.26	...	9.49364
32	Kota-Malang	46.5633	55.0411	78.96	...	12.155
33	Kota- Probolinggo	7.2606	32.0165	70.49	...	9.35005
34	Kota-Pasuruan	5.3458	27.6511	73.23	...	7.56507
35	Kota-Mojokerto	4.4268	35.4943	75.04	...	8.49718
36	Kota-Madiun	9.2142	52.8419	78.81	...	9.6101
37	Kota-Surabaya	365.3548	128.9219	78.87	...	13.8885
38	Kota-Batu	10.2597	51.6581	71.89	...	8.39926

Lampiran 5 Data Indikator Pembangunan Ekonomi Tahun 2015

No	Kabupaten/Kota	X ₁	X ₂	X ₃	...	X ₁₃
1	Pacitan	11.5906	21.0356	64.92	...	6.47436
2	Ponorogo	14.9128	17.1926	68.16	...	6.39153
3	Trenggalek	13.6324	19.78	67.25	...	6.89497
4	Tulungagung	28.4153	27.8254	70.07	...	7.91236
5	Blitar	26.7762	23.3772	68.13	...	7.5021
6	Kediri	30.4833	19.7061	68.91	...	6.08335
7	Malang	73.8433	29.023	66.63	...	7.59744
8	Lumajang	24.4568	23.7399	63.02	...	5.98408
9	Jember	56.377	23.4211	63.04	...	6.07816
10	Banyuwangi	60.2185	37.7758	68.08	...	7.35088
11	Bondowoso	14.4849	19.0291	63.95	...	5.84309
12	Situbondo	14.7957	22.093	64.53	...	5.92546
13	Probolinggo	25.6782	22.5149	63.83	...	5.73436
14	Pasuruan	104.2862	65.9288	65.04	...	7.31836
15	Sidoarjo	146.3723	69.1316	77.43	...	13.2774
16	Mojokerto	58.8199	54.4427	70.85	...	9.08082
17	Jombang	29.148	23.4875	69.59	...	6.74517
18	Nganjuk	19.1249	18.3593	69.9	...	6.14816
19	Madiun	13.8747	20.5216	69.39	...	7.59591
20	Magetan	13.8759	22.1165	71.39	...	7.11741
21	Ngawi	14.9964	18.0941	68.32	...	6.37644
22	Bojonegoro	46.6492	37.7237	66.17	...	6.42598
23	Tuban	48.2035	41.8107	65.52	...	7.43735
24	Lamongan	28.8313	24.2729	69.84	...	8.28079
25	Gresik	100.7486	80.1947	73.57	...	9.91637
26	Bangkalan	19.048	19.9602	61.49	...	5.81119

Lampiran 5 Data Indikator Pembangunan Ekonomi Tahun 2015
(Lanjutan)

No	Kabupaten/Kota	X_1	X_2	X_3	...	X_{13}
27	Sampang	14.6972	15.6888	58.18	...	5.74253
28	Pamekasan	12.3117	14.5648	63.1	...	5.32525
29	Sumenep	26.998	25.1824	62.38	...	7.18518
30	Kota-Kediri	97.4443	348.0152	75.67	...	10.1058
31	Kota-Blitar	4.8191	34.9463	76	...	11.2035
32	Kota-Malang	51.828	60.881	80.05	...	12.6019
33	Kota- Probolinggo	8.072	35.2487	71.01	...	10.1145
34	Kota-Pasuruan	5.9494	30.5412	73.78	...	10.9765
35	Kota-Mojokerto	4.8811	38.8311	75.54	...	10.9068
36	Kota-Madiun	10.1916	58.2375	79.48	...	13.8533
37	Kota-Surabaya	406.1968	142.5952	79.47	...	17.23
38	Kota-Batu	11.5104	57.4084	72.62	...	11.6584

Lampiran 6 Output Deteksi Outlier data tahun 2011

```
> rekap2011=read.csv(file.choose(),header=T, sep=";")
> data2011=data.frame(rekap2011[3:15])
> dd.plot(data2011, quan=0.9, alpha=0.05)
$outliers
 [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[25] FALSE FALSE FALSE FALSE TRUE TRUE FALSE FALSE FALSE FALSE FALSE
[37] TRUE FALSE

$md.cla
 [1] 3.085960 2.870135 2.944338 2.085624 3.010369 2.518403 2.869416 3.272148
 [9] 3.171378 2.779521 4.048036 3.220518 3.264598 3.444085 2.897377 3.060157
[17] 2.972220 2.832660 2.637815 3.284943 2.425877 3.109440 3.086591 3.759217
[25] 2.674589 4.023100 4.643676 3.305426 5.725078 5.894068 3.969175 4.188451
[33] 3.345060 4.442331 4.144761 3.204549 5.555846 3.369530

$md.rob
 [1] 2.635123 2.505593 2.667317 1.816716 2.574799 2.436198 2.963083
 [8] 2.785350 2.904945 2.571462 3.627772 2.949923 2.804918 4.050220
[15] 4.016192 2.744529 2.739858 3.227921 2.229754 2.854288 2.189165
[22] 2.926675 2.679829 3.189973 3.207676 3.556215 3.969286 3.133513
[29] 16.539454 27.786901 3.308156 3.600755 3.148210 3.934774 3.524381
[36] 3.059706 13.014468 3.976920
```

Lampiran 7 Output Deteksi Outlier data tahun 2012

```
> rekap2012=read.csv(file.choose(),header=T, sep=";")
> data2012=data.frame(rekap2012[3:15])
> dd.plot(data2012, quan=0.9, alpha=0.05)
$outliers
 [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[25] FALSE FALSE FALSE FALSE TRUE TRUE FALSE FALSE FALSE FALSE FALSE
[37] TRUE FALSE

$md.cla
 [1] 3.254405 2.274220 3.274416 3.471744 2.211384 1.732610 3.203907 4.458024
 [9] 3.930801 2.650677 4.531104 2.713209 2.843403 3.974980 2.636161 2.029808
[17] 3.608429 2.452462 3.468587 2.686034 3.696502 2.542923 3.168748 3.022325
[25] 3.241738 3.773413 3.845112 3.478481 5.907847 5.907879 3.291661 4.836946
[33] 1.866826 2.612615 4.194747 3.998447 5.549768 3.340038

$md.rob
 [1] 3.599519 1.973432 2.781368 3.042621 1.990250 1.929063 2.973054
 [8] 4.024327 3.322481 2.563402 3.918385 2.322214 2.588744 3.685871
[15] 4.294939 2.763017 3.221010 2.353478 3.124891 2.364958 3.296205
[22] 2.285489 3.192066 2.750820 3.527390 3.345409 3.351953 3.124354
[29] 29.706718 29.771632 2.924304 4.087527 1.951812 2.378022 3.638297
[36] 3.695735 12.538068 3.479493
```

Lampiran 8 Output Deteksi Outlier data tahun 2013

```
> rekap2013=read.csv(file.choose(),header=T, sep=";")
> data2013=data.frame(rekap2013[3:15])
> dd.plot(data2013, quan=0.9, alpha=0.05)
$outliers
 [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[25] FALSE FALSE FALSE FALSE TRUE TRUE FALSE FALSE FALSE FALSE FALSE FALSE
[37] TRUE FALSE

$md.cla
 [1] 4.265641 2.185005 2.819462 2.400286 3.775783 2.019262 3.650661 3.548362
 [9] 3.640244 3.112997 3.884624 2.079041 4.062538 2.764606 2.874530 2.325125
[17] 3.632759 3.542351 2.562845 2.667780 2.892803 3.188865 2.252884 2.282816
[25] 3.043833 3.847587 4.117393 3.238630 5.831888 5.922710 3.917929 3.700300
[33] 4.907482 2.185890 3.447918 4.286035 5.444997 3.637527

$md.rob
 [1] 3.640328 1.935074 2.519144 2.164704 3.187707 1.723696 3.424414
 [8] 3.199820 3.605653 2.828276 3.425800 2.780460 3.422766 3.503818
[15] 4.167335 2.220094 3.509674 2.968740 2.279688 2.268889 2.523666
[22] 2.872618 2.245961 2.317152 3.665088 3.309231 3.448640 2.769281
[29] 21.622982 32.717101 3.817482 3.260133 4.207903 2.197637 3.178755
[36] 3.717009 11.971308 3.572689
```

Lampiran 9 Output Deteksi Outlier data tahun 2014

```
> rekap2014=read.csv(file.choose(),header=T, sep=";")
> data2014=data.frame(rekap2014[3:15])
> dd.plot(data2014, quan=0.9, alpha=0.05)
$outliers
 [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[25] FALSE FALSE FALSE FALSE TRUE TRUE FALSE FALSE FALSE FALSE FALSE
[37] TRUE FALSE

$md.cla
 [1] 3.574996 2.710929 3.419876 2.887237 3.243538 2.362406 2.258096 3.355351
 [9] 3.317727 3.425750 3.378495 1.885929 3.613582 3.056856 3.465779 1.669190
[17] 2.644771 4.172396 2.275850 4.388050 3.219072 2.951701 3.222682 2.493764
[25] 3.511606 4.984986 4.094722 3.296209 5.939982 5.919095 2.234720 3.806913
[33] 4.088082 3.081806 2.763719 4.080489 5.550495 3.608673

$md.rob
 [1] 3.048967 2.380947 3.120132 2.413134 2.863755 2.451437 2.917519
 [8] 3.223255 3.653647 2.998348 3.444101 1.878390 3.394381 3.604876
[15] 4.203314 1.746047 2.479915 3.485006 2.136567 3.721543 2.765707
[22] 2.947921 2.813099 2.347191 3.643176 4.161918 3.584395 3.064925
[29] 36.036738 32.134216 2.014011 3.532232 3.459789 2.709020 2.549979
[36] 3.873412 12.974902 3.392016
```

Lampiran 10 Output Deteksi Outlier data tahun 2015

```
> rekap2015=read.csv(file.choose(),header=T, sep=";")
> data2015=data.frame(rekap2015[3:15])
> dd.plot(data2015, quan=0.9, alpha=0.05)
$outliers
 [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[25] FALSE FALSE FALSE FALSE TRUE TRUE FALSE FALSE FALSE FALSE FALSE
[37] TRUE FALSE

$md.cla
 [1] 4.129243 1.902923 3.081598 1.663152 2.824519 4.524790 2.472896 3.706578
 [9] 3.652355 2.809062 2.998477 4.504361 3.337510 3.553232 2.520108 1.862727
[17] 2.772961 4.226473 3.566820 3.109403 2.895473 2.742682 2.736835 2.054423
[25] 3.317208 4.226916 4.356370 4.501746 5.844169 5.926564 2.901475 4.709649
[33] 3.459927 2.243717 1.977282 3.460597 5.477097 3.036954

$md.rob
 [1] 3.495981 1.704159 3.245013 1.416976 2.395250 3.838817 2.852702
 [8] 3.222205 3.149822 2.506628 2.731689 3.790794 3.206619 3.689732
[15] 4.368532 2.138378 2.528581 3.607947 3.180261 2.811192 2.604952
[22] 2.543107 2.532422 1.922992 3.972049 3.621901 3.797601 3.813086
[29] 22.854661 34.183457 2.555953 3.996160 2.940787 3.329885 1.882257
[36] 3.006957 12.834806 3.006582
```

Lampiran 11 Tata Nama *Model-Based Clustering* dalam *Software R Package Teigen*

No	Model	λ_g	D_g	A_g	D_g^T
1	UUUU	U	U	U	U
2	UUUC	U	U	U	C
3	CUCU	C	U	C	U
4	CUCC	C	U	C	C
5	CUUU	C	U	U	U
6	CUUC	C	U	U	C
7	CCCU	C	C	C	U
8	CCCC	C	C	C	C
9	CIUU	C	I	U	U
10	CIUC	C	I	U	C
11	CICU	C	I	C	U
12	CICC	C	I	C	C
13	UIIU	U	I	I	U
14	UIIC	U	I	I	C
15	CIIU	C	I	I	U
16	CIIC	C	I	I	C
17	UIUU	U	I	U	U
18	UIUC	U	I	U	C
19	UCCU	U	C	C	U
20	UCCC	U	C	C	C
21	UUCU	U	U	C	U
22	UUCC	U	U	C	C
23	UICU	U	I	C	U
24	UICC	U	I	C	C
25	UCUU	U	C	U	U
26	UCUC	U	C	U	C
27	CCUU	C	C	U	U
28	CCUC	C	C	U	C

Keterangan :

C : terbatas (*constrained*), artinya antar *cluster* sama

U : tidak terbatas (*unconstrained*), artinya antar *cluster* berbeda

I : matriks Identitas

Lampiran 14 Output nilai ICL data tahun 2013

Lampiran 17 Output klasifikasi

2011

```
$classification
[1] 2 3 2 3 3 3 2 2 2 3 2 2 2 1 3 3 3 3 3 2 2 3 3 2 2 2 1 1 1 3 1 1 1 1 1 1

$modelname
[1] "CICC"

$G
[1] 3
```

2012

```
$classification
[1] 3 2 2 2 2 2 3 3 3 2 3 3 3 3 1 2 2 2 2 2 2 3 2 2 1 3 3 3 3 1 1 1 2 1 1 1 1 1 1

$modelname
[1] "CICU"

$G
[1] 3
```

2013

```
$classification
[1] 2 1 2 1 2 1 2 2 2 1 2 2 2 2 3 1 1 1 1 1 1 1 2 1 1 2 2 2 1 3 3 3 1 1 3 3 3 1

$modelname
[1] "CICU"

$G
[1] 3
```

2014

```
$classification
[1] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 1 2 2 2 2 2 2 2 2 1 2 2 2 2 1 1 1 1 1 1 1 1 1 1

$modelname
[1] "CIUC"

$G
[1] 2
```

2015

```
$classification
[1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 1 1 2 1 1 1 1 2 2 2 2 2 2 2 2 2 2

$modelname
[1] "CIUC"

$G
[1] 2
```

Lampiran 18 Output perbandingan *mean*

var	clas	Tahun				
		2011	2012	2013	2014	2015
X ₁	1	63.1846	69.3816	27.2662	68.2597	30.2965
	2	19.0626	19.9926	21.6436	28.0026	75.5695
	3	20.3955	20.6020	88.9995		
X ₂	1	53.17796	61.2250	25.2853	76.6030	25.1625
	2	23.2867	19.8949	26.9614	23.4473	84.2061
	3	20.9971	26.2368	83.5114		
X ₃	1	73.9544	74.4580	68.5447	74.7508	66.1858
	2	60.8833	67.1667	62.9407	65.4292	75.4558
	3	67.0543	61.4846	76.4814		
X ₄	1	98.4244	99.3230	99.2082	99.5792	99.2881
	2	97.9427	98.7660	98.6800	99.1988	99.9542
	3	98.8721	98.1077	99.1257		
X ₅	1	96.4656	96.0490	95.8147	98.8158	96.0892
	2	86.5573	94.7560	89.1307	95.9754	98.9992
	3	93.0121	87.5415	96.4800		
X ₆	1	74.6544	75.1520	69.8535	78.0017	68.8050
	2	50.3360	68.1927	53.8707	69.2600	83.1925
	3	63.6321	52.5269	72.9914		
X ₇	1	99.3233	99.7330	98.6806	99.7983	98.5962
	2	96.2760	99.6180	98.9364	98.5627	99.5083
	3	98.6836	97.2946	99.0871		
X ₈	1	4.9689	5.7860	4.2212	5.3492	3.8154
	2	3.7013	3.9833	3.6414	3.7065	5.5350
	3	4.0786	3.3177	6.2429		
X ₉	1	68.3956	66.3690	70.1347	66.7133	68.9792
	2	69.7213	70.2647	71.4107	69.4685	66.7375
	3	69.5907	72.1469	67.3271		
X ₁₀	1	52.9622	56.8380	65.7688	59.9317	70.6062
	2	70.3327	71.3527	69.9864	68.0004	63.0658
	3	65.1521	61.1531	58.1371		
x11	1	29.9078	28.8440	27.5176	31.7033	33.8331
	2	26.6340	27.8440	26.7800	30.9431	36.7225
	3	29.9007	26.8046	30.9357		

Lampiran 18 Output perbandingan *mean*(Lanjutan)

var	clas	Tahun				
		2011	2012	2013	2014	2015
X ₁₂	1	6.7200	7.1360	12.2329	7.3033	14.3854
	2	18.2920	12.8400	15.9743	14.3085	7.3550
	3	13.5807	17.4915	6.4157		
X ₁₃	1	7.1682	7.4548	5.7954	9.3049	6.6647
	2	3.6988	4.7623	4.6094	5.5257	11.7438
	3	4.5454	3.9530	8.7139		

Lampiran 19 Surat Legalitas Data**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini, mahasiswa jurusan Statistika FMIPA ITS:

Nama : Shiddiqul Akhyar

NRP : 1313100068

menyatakan bahwa data yang saya gunakan dalam Tugas Akhir/Thesis ini merupakan data sekunder yang diambil dari penelitian/buku/Tugas Akhir/Thesis/publikasi lainnya yaitu:

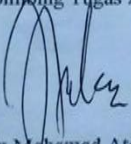
Sumber : Buku Laporan Tahunan Statistik Provinsi Jawa Timur Tahun 2011-2015, BPS Provinsi Jawa Timur


Keterangan : Data PDRB, IPM, APS, Penerangan Listrik, Ketenagakerjaan, Kriminalitas, Keluhan Kesehatan, Kemiskinan dan Pengeluaran Perkapita Kabupaten/Kota di Jawa Timur Tahun 2011-2015

Surat pernyataan ini dibuat dengan sebenar-benarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui,
Pembimbing Tugas Akhir

Surabaya, Juli 2017


Raden Mohamad Atok, Ph.D
NIP. 19710915 199702 1 001


Shiddiqul Akhyar
NRP. 1313100068

(Halaman ini sengaja dikosongkan)